

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 20:28:20 : Search time 84 Seconds
(without alignments)
22.675 Million cell updates/sec

Title: US-09-995-804A-2

Perfect score: 81

Sequence: 1 THRPPWSPVMP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum match 100%

Listing first 45 summaries

Database : A Geneseq 19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	12	23	Peptide #2 capable
2	53	65.4	86	22	Human polypeptide
3	50.5	62.3	119	22	Novel human diago
4	50	61.7	127	22	Human liver peptid
5	50	61.7	127	22	Peptide #541 enco
6	50	61.7	127	22	Protein #5266 enco
7	50	61.7	127	22	Human brain expres
8	50	61.7	127	22	Human bone marrow
9	50	61.7	127	22	Peptide #5343 enco

10	50	61.7	127	22	AA031452	Peptide #5489 enco
11	50	61.7	127	23	ABG40968	Human peptide enco
12	48	59.3	127	21	AB040488	Human ORFX ORF232
13	48	59.3	127	23	ABP06024	Human ORFX protein
14	48	59.3	133	22	AAU54613	Protonibacterium
15	48	59.3	133	22	AAU56637	Protonibacterium
16	47	58.0	97	22	AAU5736	Protonibacterium
17	47	58.0	127	22	AAU77289	Human colon cancer
18	47	58.0	1938	24	ABP76681	Streptomyces virid
19	46	56.8	55	22	AAU61586	Protonibacterium
20	46	56.8	72	22	AAU45879	Protonibacterium
21	46	56.8	75	23	ABP00471	Human ORFX protein
22	46	56.8	132	22	AAU78866	Human zinc finger
23	45.5	56.2	116	22	AAU40996	Protonibacterium
24	45.5	56.2	1938	24	ABP76680	Streptomyces virid
25	45	55.6	94	22	ABG06211	Novel human diago
26	45	55.6	157	22	ABG13699	Novel human diago
27	45	55.6	278	22	AAU87395	Novel central nerv
28	45	55.6	350	22	ABG19257	Novel human diago
29	44.5	54.9	134	22	AAU51720	Protonibacterium
30	44.5	54.9	876	20	AAU04883	Mycobacterium spec
31	44.5	54.9	886	20	AAU04884	Mycobacterium spec
32	44	54.3	20	23	AAU90618	Insulin/insulin-li
33	44	54.3	20	23	AAU90940	Insulin/insulin-li
34	44	54.3	51	22	AB04417	Peptide #11923 enc
35	44	54.3	51	22	AB027265	Protein #9264 enco
36	44	54.3	51	22	AAU78159	Human bone marrow
37	44	54.3	51	22	AAU22031	Peptide #8465 enco
38	44	54.3	51	22	AAU38403	Peptide #12440 enc
39	44	54.3	51	23	ABG47172	Human peptide enco
40	44	54.3	62	21	AAU39417	Human secreted pro
41	44	54.3	249	22	ABG17582	Novel human diago
42	44	54.3	595	22	ABG16868	Novel human diago
43	44	54.3	1001	22	ABG17352	Novel human diago
44	44	54.3	1195	21	AAU84751	Amino acid sequenc
45	44	54.3	1196	21	AAU97834	Murine circadian r

ALIGNMENTS

RESULT 1

AAU98658

ID AAU98658 standard; peptide; 12 AA.

XX AAU98658;

AC AAU98658;

DT 24-SEP-2002 (first entry)

XX Peptide #2 capable of binding to and internalising htfR.

DE Peptide #2 capable of binding to and internalising htfR.

XX Human transferrin receptor; htfR; malignancy; oral cancer;

KW liver cancer; pancreatic cancer; prostate cancer; cytostatic.

XX Synthetic.

OS Synthetic.

XX WO200244329-A2.

PN WO200244329-A2.

XX 06-JUN-2002.

PD 06-JUN-2002.

XX 29-NOV-2001; 2001WO-US44518.

PF 29-NOV-2001; 2001WO-US44518.

XX 30-NOV-2000; 2000US-253940P.

PR 30-NOV-2000; 2000US-253940P.

XX (UABR-) UAB RES FOUND.

PA (UABR-) UAB RES FOUND.

XX Engler J, Lee JH, Collawan J, Moore B;

PI Engler J, Lee JH, Collawan J, Moore B;

XX WPI; 2002-508504/54.

DR WPI; 2002-508504/54.

XX Novel peptide useful for targeting other peptides and proteins into

PT cells expressing human transferrin receptor -

XX cells expressing human transferrin receptor -

PS Claim 1; Page 19; 26pp; English.

XX The present invention relates to peptides capable of binding to and

CC internalising with the human transferrin receptor (hTfR). The peptides

CC of the invention are useful for targeting other peptides and proteins

CC into cells expressing human transferrin receptor (hTfR), and also for

CC facilitating entry of diagnostically and therapeutically useful agents,

CC including peptides and proteins. They are particularly useful for the

CC study and treatment of malignancies in cancers (e.g. oral, liver,

CC pancreatic, and prostate cancers). The present sequence represents a

CC peptide that is capable of binding to and internalising hTfR.

XX

XX Sequence 12 AA;

Query Match 100.0%; Score 81; DB 23; Length 12;

Best Local Similarity 100.0%; Pred. No. 4.4e-05;

Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 THRPPMWSVPWP 12

Db 1 THRPPMWSVPWP 12

|||||

|||||

RESULT 2

AA001416

ID RAO01416 standard; Protein; 86 AA.

XX

AC AAO01416;

XX

XX 06-NOV-2001 (first entry)

XX

XX Human polypeptide SEQ ID NO 15308.

XX

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

KW nervous system disorders; arthritis; inflammation.

XX

XX Homo sapiens.

OS

XX WO200164835-A2.

PN

XX 07-SEP-2001.

XX

XX 26-FEB-2001; 2001WO-US04927.

XX

XX 28-FEB-2000; 2000US-051526.

PR

PR 18-MAY-2000; 2000US-0577409.

XX

XX (HYSE-) HYSEQ INC.

PA

PT Tang YT, Liu C, Drmanac RT;

XX

XX WPI; 2001-514838/56.

DR

DR N-PSDB; AA181347.

XX

XX Isolated nucleic acids and polypeptides, useful for preventing

PT diagnosing and treating e.g. leukaemia, inflammation and immune

PT disorders -

XX

XX Claim 20; SEQ ID NO 15308; 1399pp + Sequence Listing; English.

XX

XX The invention relates to human polynucleotides (AA179941-AA193841) and

CC the encoded proteins (AA000010-AA013910) that exhibit activity relating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation.

CC

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX

XX Sequence 86 AA;

Query Match 65.4%; Score 53; DB 22; Length 86;

Best Local Similarity 77.8%; Pred. No. 2.3;

Matches 7; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 PPMWSVPWP 12

Db 60 PPMWSVPWP 68

|||||

|||||

RESULT 3

ABG23503

ID ABG23503 standard; Protein; 119 AA.

XX

XX ABG23503;

AC

XX 18-FEB-2002 (first entry)

DT

XX

XX Novel human diagnostic protein #23494.

DE

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder.

KW

XX Homo sapiens.

OS

XX WO200175067-A2.

PN

XX 11-OCT-2001.

PC

XX 30-MAR-2001; 2001WO-US08631.

PF

XX 31-MAR-2000; 2000US-0540217.

PR

PR 23-JUG-2000; 2000US-0649167.

XX

XX (HYSE-) HYSEQ INC.

PA

PI Drmanac RT, Liu C, Tang YT;

XX

XX WPI; 2001-639362/73.

DR

DR N-PSDB; AAS87690.

XX

XX New isolated polynucleotide and encoded polypeptides, useful in

PT diagnostics, forensics, gene mapping, identification of mutations

PT responsible for genetic disorders or other traits and to assess

PT biodiversity -

XX

XX Claim 20; SEQ ID No 53862; 103pp; English.

PS

XX The invention relates to isolated polynucleotide (I) and

CC polypeptide (II) sequences. (I) is useful as hybridisation probes,

CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome

CC and gene mapping, and in recombinant production of (II). The

CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques

CC to restore normal activity of (II) or to treat disease states involving

CC (II). (II) is useful for generating antibodies against it, detecting or

CC quantitating a polypeptide in tissue as molecular weight markers and as

CC a food supplement. (II) and its binding partners are useful in medical

CC imaging of sites expressing (II). (I) and (II) are useful for treating

CC disorders involving aberrant protein expression or biological activity.

CC The polypeptide and polynucleotide sequences have applications in

CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity

CC and to produce other types of data and products dependent on DNA and

CC amino acid sequences. ABG00010-ABG030177 represent novel human

CC diagnostic amino acid sequences of the invention.

CC Note: The sequence data for this patent did not appear in the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 119 AA;

SQ Query Match 62.3%; Score 50.5; DB 22; Length 119;

Best Local Similarity 66.7%; Pred. No. 6.8;

Matches 8; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1 THRPPMKSVPWP 12

DB 43 TRSPPLWTPAW 53

RESULT 4

ABG52883
ID ABG52883 standard; Peptide; 127 AA.

XX AC

XX ABG52883;

XX 25-FEB-2003 (first entry)

XX Human liver peptide, SEQ ID No 31531.

XX Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;

KW hypercholesterolaemia; coronary heart disease.

XX OS

XX Homo sapiens.

XX WO200157273-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00664.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GS-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for

PT analysing gene expression in human adult liver -

XX Claim 27; SEQ ID No 31531; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for

CC measuring human gene expression in a sample derived from human adult

CC liver, comprising one of 13109 defined nucleotide sequences given in the

CC specification (or complements/ fragments). The probe hybridises at high

CC stringency to a nucleic acid molecule expressed in the human adult

CC liver. (I) may be used for predicting, measuring and displaying gene

CC expression in samples derived from human adult liver. The genes

CC identified may be involved in genetic liver diseases such as cirrhosis,

CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which

CC is associated with coronary heart disease. ABG47348-ABG59930 represent

CC human liver single exon encoded peptides of the invention.

CC Note: The sequence information for this patent does not appear in the

CC printed specification but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 127 AA;

XX Query Match 61.7%; Score 50; DB 22; Length 127;

XX Best Local Similarity 54.5%; Pred. No. 8.5;

XX Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 THRPPMKSVPWP 11

DB 30 TRSPPLWTPAW 40

RESULT 5

ABG38035
ID ABG38035 standard; Peptide; 127 AA.

XX AC

XX ABG38035;

XX 04-FEB-2002 (first entry)

XX Peptide #5541 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

XX OS

XX Homo sapiens.

XX WO200157277-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00669.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GS-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

XX Human genome-derived single exon nucleic acid probes useful for

PT analysing gene expression in human fetal liver -

XX Claim 27; SEQ ID NO 30670; 639pp + sequence listing; English.

XX The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC fetal liver. The present sequence is a peptide encoded by a single exon

CC nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 127 AA;

XX Query Match 61.7%; Score 50; DB 22; Length 127;

XX Best Local Similarity 54.5%; Pred. No. 8.5;

XX Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 THRPPMKSVPWP 11

DB 30 TRSPPLWTPAW 40

RESULT 6

ABG23267
ID ABG23267 standard; Protein; 127 AA.

XX AC

XX ABG23267;

XX 23-JAN-2002 (first entry)

XX DE Protein #5266 encoded by probe for measuring heart cell gene expression.

XX KW Human; gene expression; heart; microarray; vascular system;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

KW congenital heart disease.

XX OS Homo sapiens.

XX PN WO200157274-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00666.

XX PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WI; 2001-488699/53.

XX PT Single exon nucleic acid probes for analyzing gene expression in human hearts -

XX C:aim 15; SEQ ID NO 25037; 530pp; English.

XX CC The present invention relates to single exon nucleic acid probes for measuring human gene expression in a sample derived from human heart (see ABA21535-ABA41305). The present sequence is a protein encoded by one such probe. The probes may be used for predicting, measuring and displaying gene expression in samples derived from the human heart via microarrays. By measuring gene expression, the probes are useful for predicting, diagnosing, grading, staging, monitoring and prognosing diseases of the human heart and vascular system e.g. cardiovascular disease, hypertension, cardiac arrhythmias and congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pat_sequences.

XX SQ Sequence 127 AA;

Query Match: 61.7%; Score 50; DB 22; Length 127;

Best Local Similarity 54.5%; Pred No. 8.5;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 THRPPMSPVW 11
| | | | |

Db 30 TRSPPLWTPAW 40

RESULT 7

AAM58665

ID AAM58665 standard; Protein; 127 AA.

XX AC AAM58665;

XX DT 05-NOV-2001 (first entry)

XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30770.

XX KW Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

KW epilepsy; cancer.

XX OS Homo sapiens.

XX PN WO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00667.

XX PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WI; 2001-483446/52.

XX PT Single exon nucleic acid probes for analyzing gene expression in human brains -

XX Example 4; SEQ ID NO: 30770; 650pp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid probes which are derived from genomic sequences expressed in the human brain. They can be used to measure gene expression in brain cell samples, which may enable the diagnosis and improved treatment of nervous system diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, epilepsy and cancers. The present sequence is a protein encoded by one of the probes of the invention.

XX SQ Sequence 127 AA;

Query Match: 61.7%; Score 50; DB 22; Length 127;

Best Local Similarity 54.5%; Pred No. 8.5;

Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 THRPPMSPVW 11
| | | | |

Db 30 TRSPPLWTPAW 40

RESULT 8

AAM71171

ID AAM71171 standard; Protein; 127 AA.

XX AC AAM71171;

XX DT 06-NOV-2001 (first entry)

XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31477.

XX KW Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX OS Homo sapiens.

XX PN WO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US00668.

XX PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2001-488900/53.
 XX PT Human genome-derived single exon nucleic acid probes useful for
 XX PT analyzing gene expression in human bone marrow -
 XX PS Example 4; SEQ ID NO: 31477; 658pp + Sequence Listing; English.
 XX CC The present invention provides a number of single exon nucleic acid
 XX CC probes which are derived from genomic sequences expressed in the human
 XX CC bone marrow. They can be used to measure gene expression in bone marrow
 XX CC samples, which may enable the improved diagnosis and treatment of cancers
 XX CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 XX CC protein encoded by one of the probes of the invention.
 XX SQ Sequence 127 AA;
 XX
 XX Query Match 61.7%; Score 50; DB 22; Length 127;
 XX Best Local Similarity 54.5%; Pred. No. 8.5;
 XX Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 THRPPMSPVW 11
 DB 30 TRSPPLWTPAW 40
 XX
 XX RESULT 9
 XX AAM18909
 XX ID AAM18909 standard; Protein; 127 AA.
 XX AC AAM18909;
 XX DT 12-OCT-2001 (first entry)
 XX DE Peptide #5343 encoded by probe for measuring cervical gene expression.
 XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
 XX KW cervical cancer.
 XX OS Homo sapiens.
 XX PN WO200157278-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00670.
 XX PR 04-FEB-2000; 2000US-C180312.
 XX PR 26-MAY-2000; 2000US-C207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-C234687.
 XX PR 27-SEP-2000; 2000US-C236359.
 XX PR 04-OCT-2000; 2000GB-C024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2001-488901/53.
 XX PT Human genome-derived single exon nucleic acid probes useful for
 XX PT analyzing gene expression in human cervical epithelial cells -
 XX PS Claim 27; SEQ ID NO 23735; 487pp; English.
 XX CC The present invention relates to human single exon nucleic acid probes
 XX CC (SENP: see A110068-AA28459). The present sequence is a peptide encoded
 XX CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs

CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX SQ Sequence 127 AA;
 XX
 XX Query Match 61.7%; Score 50; DB 22; Length 127;
 XX Best Local Similarity 54.5%; Pred. No. 8.5;
 XX Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 THRPPMSPVW 11
 DB 30 TRSPPLWTPAW 40
 XX
 XX RESULT 10
 XX AAM31452
 XX ID AAM31452 standard; Protein; 127 AA.
 XX AC AAM31452;
 XX DT 17-OCT-2001 (first entry)
 XX DE Peptide #5489 encoded by probe for measuring placental gene expression.
 XX KW Probe; microarray; human; placenta; antenatal diagnosis;
 XX KW genetic disorder.
 XX OS Homo sapiens.
 XX PN WO200157272-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US00663.
 XX PR 04-FEB-2000; 2000US-0180312.
 XX PR 26-MAY-2000; 2000US-0207456.
 XX PR 30-JUN-2000; 2000US-0608408.
 XX PR 03-AUG-2000; 2000US-0632366.
 XX PR 21-SEP-2000; 2000US-0234687.
 XX PR 27-SEP-2000; 2000US-0236359.
 XX PR 04-OCT-2000; 2000GB-0024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX DR WPI; 2001-488997/53.
 XX PT Human genome-derived single exon nucleic acid probes useful for
 XX PT analyzing gene expression in human placenta -
 XX PS Claim 27; SEQ ID No 31721; 654pp; English.
 XX CC The present invention relates to single exon nucleic acid probes (SENP:
 XX CC see A113115-AA157546). The present sequence is a peptide encoded by one
 XX CC such probe. The probes are useful for producing a microarray for
 XX CC predicting, measuring and displaying gene expression in samples derived
 XX CC from human placenta. The probes are useful for antenatal diagnosis of
 XX CC human genetic disorders.
 XX SQ Sequence 127 AA;
 XX
 XX Query Match 61.7%; Score 50; DB 22; Length 127;
 XX Best Local Similarity 54.5%; Pred. No. 8.5;
 XX Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 THRPPKMSPPW 11
| | | | |
Db 30 TRSPPLWTPAW 40

RESULT 11
ABG40968
ID ABG40968 standard; Peptide: 127 AA.
XX AC ABG40968;
XX 19-AUG-2002 (first entry)
XX Human peptide encoded by genome-derived single exon probe SEQ ID 30633.
KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.
XX Homo sapiens.
XX WO2001:86003-A2.
XX 15-NOV-2001.
XX 30-JAN-2001; 2001WO-US00665.
XX 04-FEB-2000; 2000US-180312P.
XX 26-MAY-2000; 2000US-207456P.
XX 30-JUN-2000; 2000US-0608468.
XX 03-AUG-2000; 2000US-0632366.
XX 21-SEP-2000; 2000US-2346872.
XX 27-SEP-2000; 2000US-236359P.
XX 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples -
XX Claim 27; SEQ ID No 30633; 634pp; English.
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of
XX probes; the novel set of probes which hybridise at high stringency to a
XX nucleic acid expressed in the human lung; measuring gene expression in a
XX sample derived from human lung, comprising (a) contacting the array with
XX a collection of detectably labeled nucleic acids derived from human lung
XX mRNA, and (b) measuring the label detectably bound to each probe of
XX the array; identifying exons in a eukaryotic genome, comprising
XX (a) algorithmically predicting at least one exon from genomic sequences
XX of the eukaryote; and (b) detecting specific hybridisation of detectably
XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
XX having a fragment identical to the predicted exon, the probe is included
XX in the above mentioned microarray; assigning exons to a single gene,
XX comprising (a) identifying exons from genomic sequence by the method
XX above, and (b) measuring the expression of each of the exons in several
XX tissues and/or cell types using hybridisation to a single exon
XX microarrays, having a probe with the exon, where a common pattern of

CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
CC and hyaline membrane disease. The present sequence is a peptide/protein
CC encoded by a single exon probe of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 127 AA;
Query Match 61.7%; Score 50; DB 23; Length 127;
Best Local Similarity 54.5%; Pred. No 8.5;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 1 THRPPKMSPPW 11
| | | | |
Db 30 TRSPPLWTPAW 40

RESULT 12
AAB40468
ID AAB40468 standard; Protein: 127 AA.
XX AC AAB40468;
XX 08-FEB-2001 (first entry)
XX Human ORF332 polypeptide sequence SEQ ID NO:464.
XX Human; open reading frame; ORF3; detection; cytostatic; hepatotropic;
XX vulnary; antipapillary; antipapillary; antipapillary; antipapillary;
XX anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiac;
XX immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
XX hypotensive; dermatologic; immunosuppressive; antiinflammatory;
XX antiviral; antibacterial; antifungal; antineoplastic; antichyroid;
XX antianemic; gene therapy; cancer; proliferative disorder; hypertension;
XX neurodegenerative disorder; osteoarthritis; graft vs host disease;
XX cardiovascular disease; diabetes mellitus; hypothyroidism; SCD; AIDS;
XX cholesterol ester storage; systemic lupus erythematosus; infection;
XX severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
XX allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
XX bone damage; cartilage damage; antiinflammatory disease; coagulation;
XX thrombosis; contraceptive.
XX Homo sapiens.
XX WO2000058473-A2.
XX 05-OCT-2000.
XX 31-MAR-2000; 2000WO-US08621.
XX 31-MAR-1999; 99US-0127607.
XX 02-APR-1999; 99US-0127636.
XX 05-APR-1999; 99US-0127928.
XX 30-MAR-2000; 2000US-0540763.
XX (CURA-) CURAGEN CORP.
XX Shimkets RA, Leach M;
XX

DR WP*: 2000-602362/57.
 DR N-PSDB; AAC74677.
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 PS Claim 11; Page 661-662; 5507pp; English.
 XX
 CC AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 CC which represent the human ORFX open reading frames 1 to 3161. The ORFX
 CC sequences have activities such as: cytostatic; hepatotropic; vulnary;
 CC antiproliferative; antiparkinsonian; nocrotropic; neuroprotective;
 CC osteopathic; aniconvulsant; antiarthritic; immunosuppressant;
 CC immunostimulant; cardiac; thrombolytic; coagulant; vasotropic;
 CC antidiabetic; hypotensive; dermatological; immunosuppressive;
 CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;
 CC antihypertensive; and antianemic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORFX-associated disorder. The
 CC nucleic acids can be used to express ORFX proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.
 XX
 SQ Sequence 127 AA;
 Query Match 59.3%; Score 48; DB 2; Length 127;
 Best Local Similarity 58.3%; Pred. No. 16;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 THRPMMSPVWP 12
 ||| :||| :|
 Db 41 THHPALWSPDFP 52

RESULT 13
 ABP06024
 ID ABP06024 standard; Protein; 127 AA.
 AC ABP06024;
 XX
 DT 24-JUN-2002 (first entry)
 DE Human ORFX protein sequence SEQ ID NO:12030.
 XX
 KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.
 XX
 OS Homo sapiens.
 XX
 PN WO200192523-A2.
 XX
 PD 06-DEC-2001.
 XX
 XX 29-MAY-2001; 2001WO-US10836.
 PF
 XX 30-MAY-2000; 2000US-206132P.
 PR
 XX 29-AUG-2000; 2000US-228716P.
 XX
 PA (CURA-) CURAGEN CORP.

XX Shimketa RA, Leach MD;
 XX
 DR WPI; 2002-106308/14.
 DR N-PSDB; ABB21776.
 XX
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders -
 XX
 PS Disclosure; SEQ ID 12030; 1037pp; English.
 XX
 CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABB21776 to ABB21782 encode the human ORFX
 CC proteins given in ABB0010 to ABB1500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC diseases, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 127 AA;
 Query Match 59.3%; Score 48; DB 23; Length 127;
 Best Local Similarity 58.3%; Pred. No. 16;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 THRPMMSPVWP 12
 ||| :||| :|
 Db 41 THHPALWSPDFP 52

RESULT 14
 AAU54613
 ID AAU54613 standard; Protein; 133 AA.
 AC AAU54613;
 XX
 DT 27-FEB-2002 (first entry)
 DE Propionibacterium acnes immunogenic protein #15509.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 XX 20-APR-2001; 2001WO-US12865.
 PF
 XX 21-APR-2000; 2000US-199047P.
 PR

```

PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
XX N-PSDB; AAS59566.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 13808; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA).
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 133 AA;

Query Match 59.3%; Score 48; DB 22; Length 133;
Best Local Similarity 58.3%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 THRPWMSPVWP 12
   ||| :||| :|
DB 82 THHPALWSPDFP 93

Search completed: October 22, 2003, 20:38:18
Job time : 87 secs

```

```

PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX WPI; 2001-616774/71.
XX N-PSDB; AAS59566.
XX
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 13808; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA).
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 133 AA;

Query Match 59.3%; Score 48; DB 22; Length 133;
Best Local Similarity 58.3%; Pred. No. 17;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 THRPWMSPVWP 12
   ||| :||| :|
DB 82 THHPALWSPDFP 93

RESULT 15
ID AAU5637 standard; Protein; 133 AA.
XX
XX AAU5637;
XX AC
XX
XX 27-FEB-2002 (first entry)
XX
XX Propionibacterium acnes immunogenic protein #26533.
XX
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
XX dermatological; osteopathic; neuroprotectant.
XX
XX Propionibacterium acnes.
XX OS
XX WO20018158-A2.
XX PN
XX
XX 01-NOV-2001.
XX PD
XX
XX 20-APR-2001; 2001WO-US12865.
XX PF
XX
XX 21-APR-2000; 2000US-199047P.
XX PR

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 20:36:51 ; Search time 40 seconds
(without alignments)
28.951 Million cell updates/sec

Title: US-09-995-804A-2

Perfect score: 81

Sequence: 1 THRPMPKSPVWP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76: *
1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	47	58.0	503	2	A83027
2	46	56.8	162	2	H75435
3	45.5	56.2	698	2	AC0016
4	44.5	54.9	876	2	D70971
5	44	54.3	946	2	T01460
6	44	54.3	1197	2	T13956
7	43	53.1	233	2	B6436
8	43	53.1	554	2	AC0320
9	41.5	51.2	410	2	B84401
10	41.5	51.2	437	2	B82778
11	41.5	51.2	966	2	G69189
12	41	50.6	94	2	T36982
13	41	50.6	134	2	I47010
14	41	50.6	174	1	H4FP22
15	41	50.6	174	2	A20647
16	41	50.6	184	2	H96718
17	41	50.6	336	2	S30972
18	41	50.6	399	2	T20439
19	41	50.6	416	2	JC1343
20	41	50.6	468	2	A22422
21	41	50.6	724	2	T04340
22	41	50.6	731	2	T17002
23	41	50.6	1447	2	S30918
24	40.5	50.0	195	2	A10252
25	40	49.4	56	2	PC4132
26	40	49.4	183	2	S29088
27	40	49.4	192	2	C87670
28	40	49.4	227	2	D87568
29	40	49.4	336	2	A72803

30	40	49.4	459	2	B44498
31	40	49.4	471	2	T33997
32	40	49.4	483	2	T05088
33	40	49.4	524	2	E75574
34	40	49.4	530	2	A13585
35	40	49.4	554	2	G75576
36	40	49.4	734	2	T24908
37	40	49.4	1067	2	D82436
38	40	49.4	1332	1	I48314
39	40	49.4	1468	2	A4345
40	40	49.4	1475	2	S42718
41	39.5	48.8	703	2	H86349
42	39	48.1	10	2	S39030
43	39	48.1	108	2	E72494
44	39	48.1	153	2	A55139
45	39	48.1	180	2	H95334

ALIGNMENTS

RESULT 1

A83027
hypothetical protein PA4955 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: A83027
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: A83027
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-503 <STO>
A:Cross-references: GB:AE004908; GB:AE004091; NID:g9951230; PIDN:AAG08340.1; GSPDB:GN000000000
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4955

Query Match 58.0%; Score 47; DS.2; Length 503;
Best Local Similarity 50.0%; Pred. No. 14;
Matches 7; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 3 RPPMWSP----VWP 12
DB 303 RPPLMRPAVALIWP 316

RESULT 2

H75435
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: H75435
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-162 <WHI>
A:Cross-references: GB:AE001961; GB:AE000513; NID:g6458843; PIDN:AAF10691.1; PID:g6458843
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR113
A:Map position: 1

Query Match 56.8%; Score 46; DB 2; Length 162;
 Best Local Similarity 60.0%; Pred. No. 6.1;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Cy 2 HRPPMWSFW 11
 ||| |||
 Db 63 HRPDHWPLW 72

RESULT 3
 AC0016
 4-alpha-glucanotransferase (EC 2.4.1.25) [imported] - *Yersinia pestis* (strain CO92)
 C:Species: *Yersinia pestis*
 C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
 C:Accession: AC0016
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougar, G.; Il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001
 A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A:Reference number: AB0001; MUID:21475413; PMID:11586360
 A:Accession: AC0016
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-698 <KUR>
 A:Cross-references: GB:AL590842; PIDN:CAC68989.1; PID:gl5978231; GSPDB:GN00175
 C:Genetics:
 A:Gene: *imaQ*
 C:Superfamily: 4-alpha-glucanotransferase
 C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 56.2%; Score 45.5; DB 2; Length 698;
 Best Local Similarity 66.7%; Pred. No. 32;
 Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Cy 2 HRPPMWS-PVWP 12
 ||| |||
 Db 313 HDPVWGWGPWP 324

RESULT 4
 D70971
 hypothetical protein RV3365c - *Mycobacterium tuberculosis* (strain H37RV)
 C:Species: *Mycobacterium tuberculosis*
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
 C:Accession: D70971
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.; Nature 393, 537-544, 1998
 A:Authors: Graves, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295997; PMID:9634230
 A:Accession: D70971
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-876 <COL>
 A:Cross-references: GB:AL123456; GB:AL123456; NID:g3242262; PIDN:CAA15750.1; PID:e120228
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: RV3365c

Query Match 54.9%; Score 44.5; DB 2; Length 876;
 Best Local Similarity 72.7%; Pred. No. 57;
 Matches 8; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

Cy 3 RPPMWS-PVWP 12
 ||| |||
 Db 32 RPPSWSPNWP 42

RESULT 5
 T01460

hypothetical protein T24H24.6 - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 02-Mar-2001
 C:Accession: T01460
 R:Courtney, L.; Stoneking, T.; Langston, Y.; Mead, K. submitted to the EMBL Data Library, August 1998
 A:Description: The sequence of *A. thaliana* T24H24.
 A:Reference number: Z14333
 A:Accession: T01460
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-946 <COU>
 A:Cross-references: EMBL:AF055598; NID:g3293591; PID:g3377831
 A:Experimental source: Cultivar Columbia
 C:Genetics:
 A:Map position: 4
 A:Introns: 618/2; 803/3
 A:Note: T24H24.6
 C:Superfamily: *Arabidopsis thaliana* hypothetical protein F26C24.7

Query Match 54.3%; Score 44; DB 2; Length 946;
 Best Local Similarity 87.5%; Pred. No. 73;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Cy 2 HRPPMWS P 9
 ||| |||
 Db 746 HRLPMWS P 753

RESULT 6
 T13956
 timeless protein homolog mTim1 - mouse
 C:Species: *Mus musculus* (house mouse)
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
 C:Accession: T13956
 R:Koike, N.; Hida, A.; Numano, R.; Hirose, M.; Sakaki, Y.; Tei, H. FEBS Lett. 441, 427-431, 1998
 A:Title: Identification of the mammalian homologues of the *Drosophila* timeless gene.
 A:Reference number: Z17832; MUID:99107002; PMID:9891984
 A:Accession: T13956
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1197 <KOI>
 A:Cross-references: EMBL:AB015598; NID:g4115715; PIDN:BAA36500.1; PID:g4115716
 A:Experimental source: strain BAB/c
 C:Genetics:
 A:Gene: mTim1

Query Match 54.3%; Score 44; DB 2; Length 1197;
 Best Local Similarity 66.7%; Pred. No. 92;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Cy 2 THRPWMSP 9
 ||| |||
 Db 810 SHRAPLWSP 818

RESULT 7
 B64436
 cobalamin biosynthesis protein M homolog - *Methanococcus jannaschii*
 C:Species: *Methanococcus jannaschii*
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C:Accession: B64436
 R:Sult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak, R.; Raich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, J.; rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
 A:Reference number: A64300; MUID:96337999; PMID:8688087
 A:Accession: B64436
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA

```

A;Residues: 1-233 <BLU>
A;Cross-references: GB:U67551; GB:L77117; NID:gl591728; PIDN:AAB99032.1; P.D:gl591735; I
C;Genetics:
C;Map position: REV1029822-1029121
C;Superfamily: cobalamin biosynthesis protein M

Query Match      53.1%; Score 43; DB 2; Length 233;
Best Local Similarity 75.0%; Pred. No. 24;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4 PPMWSPVW 11
Db 9 PPMCAW 16

RESULT 8
AC0320
asparagine synthase (glutamine-hydrolyzing) (EC 6.3.5.4) [imported] - Versinia pestis (S
C;Species: Versinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 27-Nov-2001
C;Accession: AC0320
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, X.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001.
A;Title: Genome sequence of Versinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AC0320
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-554 <KUR>
A;Cross-references: GB:AL590842; PIDN:CA092866.1; P.D:gl5980610; GSPDB:GN00175
C;Genetics:
A;Gene: asnA
C;Superfamily: asparagine synthase (glutamine-hydrolyzing)
C;Keywords: ligase

Query Match      53.1%; Score 43; DB 2; Length 554;
Best Local Similarity 70.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 2 HRPMPSPVW 11
Db 30 HRPDMSGVW 39

RESULT 9
B84401
mRNA 3'-end processing factor homolog [imported] - Halobacterium sp. NRC-1
C;Species: Halobacterium sp. NRC-1
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C;Accession: B84401
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Skukla, H.D.; Lasky, S.;
Leitthausser, B.; Keller, K.; Cruz, R.; Danson, V.C.; Hough, D.W.; Maddocks, D.G.; Gabil
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Ozer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.
A;Reference number: A84160; MUID:20504483; PMID:11016950
A;Accession: B84401
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-410 <STO>
A;Cross-references: GB:ASC04437; NID:gl0581905; PIDN:AAG20574.1; GSPDB:GN00138
C;Genetics:
A;Gene: ept1

Query Match      51.2%; Score 41.5; DB 2; Length 410;
Best Local Similarity 70.0%; Pred. No. 73;
Matches 7; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 1 THRPMP-WSP 9

```

```

Db 67 THRPPIHTP 76

RESULT 10
B82778
cell cycle protein XPO659 [imported] - Xylella fastidiosa (strain 945c)
C;Species: Xylella fastidiosa
C;Date: 16-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Dec-2002
C;Accession: B82778
R;Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: B82778
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-437 <SIM>
A;Cross-references: GB:AE003910; GB:AE003849; NID:gl05532; PIDN:AAB93469.1; GSPDB:GN
A;Experimental source: strain 945c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.;
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer,
as-Neto, E.; Docena, C.; El-Dorzy, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa
V.; Teubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A;Reference number: A59328
A;Contents: annotation
A;Genetics:
A;Gene: XPO659
C;Superfamily: Cell cycle protein MesJ

Query Match      51.2%; Score 41.5; DB 2; Length 437;
Best Local Similarity 46.7%; Pred. No. 77;
Matches 7; Conservative 1; Mismatches 2; Indels 5; Gaps 1;

Qy 2 HR-----PPMWSPVW 11
Db 311 HRPQPPPPDQWQPLW 325

RESULT 11
G69189
hypothetical protein MTH674 - Methanobacterium thermoautotrophicum (strain Delta H)
C;Species: Methanobacterium thermoautotrophicum
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 22-Oct-1999
C;Accession: G69189
R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, C.; Aldredge, T.;
Giu, D.; Spadafora, R.; Vicaire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiواني,
K.; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.
J. Bacteriol. 179, 7135-7155, 1997
A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu
A;Reference number: A69000; MUID:98037514; PMID:9371463
A;Accession: G69189
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-966 <MTH>
A;Cross-references: GB:AE000847; GB:AE000666; NID:gl2621756; PIDN:AAB85179.1; P.D:gl262
A;Experimental source: strain Delta H
C;Genetics:
A;Gene: MTH674

Query Match      51.2%; Score 41.5; DB 2; Length 966;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

```

QY 2 HRPPMSPVP 11
|||
Db 571 HR-PJWKPLW 579

RESULT 12

T36982
probable insertion element transposase - Streptomyces coelicolor (fragment)
C/Species: Streptomyces coelicolor
C/Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 08-Sep-2000
C/Accession: T36982
R/Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, August 1999
A/Reference number: Z21618
A/Accession: T36982
A/Status: preliminary; translated from GB/EMBL/DBSJ
A/Molecule type: DNA
A/Residues: 1-94 <OLI>
A/Cross-references: EMBL:AL109949; PIDN: CAB52896.1; GSPDB: GN00070; SCOEDB: SCJ11.11c
A/Experimental source: strain A3(2)
C/Genetics:
A/Gene: SCOEDB: SCJ11.11c

Query Match 50.6%; Score 41; DB 2; Length 94;
Best Local Similarity 55.6%; Pred. No. 19;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 PPMWSPVP 12
|||
Db 35 PPVFGPAW 43

RESULT 13

147010
gastrin-releasing peptide - sheep
C/Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C/Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 19-Jan-2001
C/Accession: I47010
R/Fraser, M.; McDonald, T.J.; Spinde, E.R.; Fahy, M.; Hill, D.; Challis, J.R.
Endocrinology 135, 2440-2445, 1994
A/Title: Gastrin-releasing peptide is produced in the pregnant ovine uterus.
A/Reference number: I47010; MUID: 95080110; PMID: 7998429
A/Accession: I47010
A/Status: preliminary; translated from GB/EMBL/DBSJ
A/Molecule type: mRNA
A/Residues: 1-134 <FRA>
A/Cross-references: GB: S75723; NID: G913167; PIDN: AAB32675.1; PID: G913168
C/Superfamily: gastrin-releasing peptide

Query Match 50.6%; Score 41; DB 2; Length 134;
Best Local Similarity 55.6%; Pred. No. 28;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HRPPMSPVP 10
|||
Db 96 HPPRWEPL 104

RESULT 14

HFF22
heat shock protein 22 - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 15-Nov-1984 #sequence_revision 15-Nov-1984 #text_change 16-Feb-1997
C/Accession: A02918
R/Ingolia, T.D.; Craig, E.A.
Proc. Natl. Acad. Sci. U.S.A. 79, 2360-2364, 1982
A/Title: Four small Drosophila heat shock proteins are related to each other and to mammalian HSP70.
A/Reference number: A93909; MUID: 82248004; PMID: 6285380
A/Accession: A02918
A/Molecule type: DNA
A/Residues: 1-174 <ING>
A/Note: the authors translated the codon CAT for residue 44 as Gln, GAG for residue 83 as Lys, and the codon usage table proposed by the authors and the translation of the protein

C/Comment: This small heat shock protein is related to alpha crystallin.

C/Genetics:
A/Gene: FlyBase: Hsp22
A/Cross-references: FlyBase: FBgn0001223
A/Map position: 3L (67B)
C/Superfamily: alpha-crystallin
C/Keywords: heat shock; stress-induced protein

Query Match 50.6%; Score 41; DB 1; Length 174;
Best Local Similarity 54.5%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 4; Indels 3; Gaps 0;

QY 2 HRPPMSPVP 12
|||
Db 29 HEPPVMSVALP 39

RESULT 15

A20647
heat shock protein 22 - fruit fly (Drosophila melanogaster)
C/Species: Drosophila melanogaster
C/Date: 09-Sep-1987 #sequence_revision 09-Sep-1987 #text_change 21-Jul-2000
C/Accession: A20647
R/Southgate, R.; Ayme, A.; Voellmy, R.
J. Mol. Biol. 165, 35-57, 1983
A/Title: Nucleotide sequence analysis of the Drosophila small heat shock gene cluster
A/Reference number: A92897; MUID: 83189140; PMID: 6302284
A/Accession: A20647
A/Molecule type: DNA
A/Residues: 1-174 <SOU>
A/Cross-references: GB: X03888; GB: J01090; GB: J01091; GB: J01098; GB: V00209; NID: G9108;

C/Genetics:
A/Gene: FlyBase: Hsp22
A/Cross-references: FlyBase: FBgn0001223
C/Superfamily: alpha-crystallin

Query Match 50.6%; Score 41; DB 2; Length 174;
Best Local Similarity 54.5%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HRPPMSPVP 12
|||
Db 29 HEPPVMSVALP 39

Search completed: October 22, 2003, 20:41:14
Job time : 43 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 20:35:06 ; Search time 23 Seconds
(without alignments)
24,536 Million cell updates/sec

Title: US-09-995-804A-2
Perfect score: 81
Sequence: 1 THRPMPMSPWP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	42.5	52.5	606	1 E75 METEN	O77245 metapanaeus
2	42	51.9	720	1 G7AC BREDI	O915d6 brevundamon
3	41	50.6	134	1 GRP SHEEP	P4785: ovis aries
4	41	50.6	174	1 HS22 DROME	P02815 drosophila
5	41	50.6	336	1 VG27 BPWL5	Q05234 mycobacteri
6	41	50.6	496	1 NMT1 MOUSE	P30419 homo sapien
7	41	50.6	496	1 NMT1 BOVIN	P17177 bos taurus
8	41	50.6	497	1 YH09 RALSO	O8xyd9 ralstonia s
9	41	50.6	657	1 BGAL_MALDO	P48981 malus domes
10	41	50.6	731	1 SGSL_YEAST	P35187 saccharomyc
11	41	50.6	1447	1 TTY7 HUMAN	Q95q92 homo sapien
12	40.5	50.0	124	1 DSB8_XANAC	O8pnd6 xanthomonas
13	40	49.4	172	1 DSB8_XANAC	O8pnc32 xanthomonas
14	40	49.4	172	1 DSB8_XANAC	O8pnc32 xanthomonas
15	40	49.4	183	1 UBCX_YEAST	P29340 saccharomyc
16	40	49.4	336	1 VG27_BPWD2	O64221 mycobacteri
17	40	49.4	459	1 RSP6_CHURE	Q01657 chlamydomon
18	40	49.4	1395	1 CUT1 MOUSE	P53564 mus musculu
19	40	49.4	1468	1 N153 RAT	P49791 rattus norv
20	40	49.4	1475	1 N153_HUMAN	P49790 homo sapien
21	39	48.1	10	1 BRK_OCNKY	O9pr21 oncorhynch
22	39	48.1	153	1 GJB2_ASCSU	P45672 ascaris suu
23	39	48.1	334	1 SCUP_CHAMU	Q9pln7 chlamydia m
24	39	48.1	780	1 YNL5_CAEEL	Q21955 caenorhabdi
25	39	48.1	887	1 GLND_KLEPN	P41393 klebsiella
26	39	48.1	1180	1 ATY1_HUMAN	O9cqi1 homo sapien
27	38.5	47.5	1257	1 PER2_MOUSE	O54943 mus musculu
28	38	46.9	365	1 OFSR_XENLA	O12948 xenopus lae
29	38	46.9	522	1 AOF_OCNKY	P49253 oncorhynch
30	38	46.9	573	1 AMH2_HUMAN	Q15671 homo sapien
31	38	46.9	585	1 F13C_HUMAN	O8re31 homo sapien
32	38	46.9	827	1 SOX6_MOUSE	P40645 mus musculu
33	38	46.9	828	1 SOX6_HUMAN	P35712 homo sapien

ALIGNMENTS

RESULT :	E75 METEN	STANDARD;	PRT;	606 AA.
ID	E75 METEN			
AC	O77245;			
DT	15-JUL-1999 (Rel. 38, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Nuclear hormone receptor E75.			
GN	E75 OR NR1D3.			
CS	Metapanaeus ensis (Greasyback shrimp) (Sand shrimp).			
OC	Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;			
OC	Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;			
OC	Penaeidae; Metapanaeus.			
OX	NCBI_TaxID=32278;			
RY	[1.._TaxID=32278;			
RN	SEQUENCE FROM N.A.			
RC	TISSUE=Pleopod;			
RX	MEDLINE=99015775; PubMed=9801156;			
RA	Chan S.-M.;			
RT	"Cloning of a shrimp (Metapanaeus ensis) cDNA encoding a nuclear			
RT	receptor superfamily member; an insect homologue of E75 gene."			
RL	FEBS Lett. 436:395-400(1998).			
CC	-!- SUBCELLULAR LOCATION: Nuclear (Potential).			
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN THE EPIDERMIS, EYESTALK AND THE			
CC	NERVE CORD OF THE PRE-MOLT SHRIMP.			
CC	-!- SIMILARITY: Belongs to the nuclear hormone receptor family. NR:			
CC	subfamily.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF092946; AAC71770.1; ..			
DR	HSSP; P20393; 1A6Y.			
DR	InterPro; IPR000536; Hormone_rec_lig.			
DR	InterPro; IPR001723; Stdrhm_receptor.			
DR	InterPro; IPR001628; Znf_C4steroid.			
DR	Pfam; PF00104; hormone_rec; 1.			
DR	Pfam; PF00105; zf-C4; 1.			
DR	PRINTS; PR00398; STRDHORMONER.			
DR	PRINTS; PR00047; STROIDFINGER.			
DR	ProDom; PD000035; Znf_C4steroid; 1.			
DR	SMART; SM00430; HOLI; 1.			
DR	SMART; SM00399; Znf_C4; 1.			
DR	PROSITE; PS00031; NUCLEAR RECEPTOR; ..			
KW	Receptor; Transcription regulation; DNA-binding; Nuclear protein;			
KW	Zinc-finger.			
FT	DNA BIND 32 98 NUCLEAR RECEPTOR-TYPE.			
FT	ZN FING 32 52 C4-TYPE.			
FT	ZN FING 69 93 C4-TYPE.			
FT	DOMAIN 166 383 LIGAND-BINDING (POTENTIAL).			
FT	SEQUENCE 606 AA; 67689 MW; 337A7FE74929438F CRC64;			

Q09092 brassica ol
P53565 rattus norv
P29020 mus musculu
P12525 homo sapien
P22083 homo sapien
Q62994 rattus norv
O43206 triticum ae
P25918 mus musculu
P24558 salmo salar
P26170 rhodobacter
Q9uyv6 pyrococcus
O55028 mus musculu

Query Match 52.5%; Score 42.5; DS 1; Length 606;
 Best Local Similarity 70.0%; Pred. No. 41;
 Matches 7; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 2 HRPMPWSPV 11
 |||||
 DB 593 HTPP-WPVPV 601

RESULT 2
 ID G7AC BREDI STANDARD; PRT; 720 AA.
 AC Q9L5D6;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Glutaryl 7-aminocephalosporanic acid acylase precursor (EC 3.5.1.1)
 DE (Glutaryl-7-ACA-acylase)
 OS Brevundimonas diminuta (Pseudomonas diminuta)
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Caulobacteriales;
 OC Caulobacteraceae; Brevundimonas.
 OX NCB: TaxID=293;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-KAC-1;
 RA Kim D.-W., Kang S.-M., Yoon K.-H.;
 RT "Cloning and the nucleotide sequence of a Pseudomonas diminuta KAC-1
 glutaryl 7-aminocephalosporanic acid acylase gene";
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- PATHWAY: SECOND STEP IN THE BIOCONVERSION OF CEPHALOSPORIN TO 7-
 ACA.
 CC -!- SUBCELLULAR LOCATION: Periplasmic (Potential).
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S45.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AF251710; AAF64242.1; -
 CC DB: JPM2; 15-AUG-01.
 CC DB: JCVZ; 01-SEP-02.
 CC DB: IOWC; 02-SEP-02.
 CC DB: IKEH; 04-DEC-02.
 CC DB: MEROPS; S45.002; -
 CC DB: InterPro: IPR002692; Penicillin amidase.
 CC DB: Pfam: PF01804; Penicillin amidase; 1.
 KW Hydrolase; Periplasmic; Antibiotic resistance; Zymogen; Signal;
 KW 3D-structure.
 FT SIGNAL 1 27 POTENTIAL.
 FT CHAIN 28 720 GLUTARYL 7-AMINOCEPHALOSPORANIC ACID
 FT ACYLASE.
 FT ACT SITE 199 199 PROBABLE.
 FT SEQUENCE 720 AA; 79779 MW; A6242797845CC39B CRC64;
 Query Match 51.9%; Score 42; DB 1; Length 720;
 Best Local Similarity 62.5%; Pred. No. 57;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 PMWSPVWP 12
 |||||
 DB 444 PPWTPWP 451

RESULT 3
 GRP_SHEEP
 ID GRP_SHEEP STANDARD; PRT; 134 AA.
 AC P47851;
 DT 01-FEB-1996 (Rel. 33, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gastrin-releasing peptide precursor (GRP) [Contains: Neuromedin C
 DE (GRP-10)].
 GN GRP.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCB: TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95080110; PubMed=7988429;
 RA Fraser M., McDonald T.J., Spindel E.R., Fahy M., Hill D.,
 RA Challis J.R.;
 RT "Gastrin-releasing peptide is produced in the pregnant ovine uterus";
 RL Endocrinology 135:2440-2445(1994).
 CC -!- FUNCTION: GRP stimulates gastrin release as well as other
 CC gastrointestinal hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: BELONGS TO THE BOMBESIN/NEUROMEDIN B/RANATENSIN
 CC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: S75723; AAB32675.1; -
 CC DB: JPIR; 147016; J47010.
 CC DB: InterPro: IPR000874; Bombesin.
 CC DB: Pfam: PF02044; Bombesin; 1.
 CC DB: PROSITE: PS00357; BOMBESIN; 1.
 KW Bombesin family; Amidation; Cleavage on pair of basic residues;
 KW Signal.
 FT SIGNAL 1 23 BY SIMILARITY
 FT PEPTIDE 24 50 GASTRIN-RELEASING PEPTIDE.
 FT PROPEP 54 134
 FT PEPTIDE 41 50 NEUROMEDIN C.
 FT MOD_RES 50 50 AMIDATION (G-51) PROVIDED AMIDE GROUP.
 SQ SEQUENCE 134 AA; 14855 MW; C8958538E4C6F0A4 CRC64;
 Query Match 50.6%; Score 41; DB 1; Length 134;
 Best Local Similarity 55.6%; Pred. No. 15;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 HRPMPWSPV 10
 |||||
 DB 96 HQPRWEPL 104

RESULT 4
 HS22 DROME STANDARD; PRT; 174 AA.
 ID HS22 DROME STANDARD; PRT; 174 AA.
 AC P02515; Q95S24; Q9VSX1;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Heat shock protein 22.
 DE HSP22 OR CG4460.
 GN
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCB: TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=83189140; PubMed=6302284;
 RA Southgate R., Ayme A., Voellmy R.;
 RT "Nucleotide sequence analysis of the Drosophila small heat shock gene

cluster at locus 67B.";
J. Mol. Biol. 165:35-57(1983).
[2]
SEQUENCE FROM N.A.
MEDLINE=8224804; PubMed=6285380;
RA Ingolia T.D., Craig E.A.;
RA "Four small Drosophila heat shock proteins are related to each other
RT and to mammalian alpha-crystallin."
Proc. Natl. Acad. Sci. U.S.A. 79:2360-2364 (1992).
[3]
SEQUENCE FROM N.A.
STRAIN=Berkley;
MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins P.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brannon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.Y., Basu A., Baxendale J., Bayraktaroglu L., Beasley B.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bozhakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.V., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck C.,
RA Hostin D., Houslan K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Labko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao X., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
Science 287:2185-2195(2000).
[4]
SEQUENCE FROM N.A.
STRAIN=Berkley; Tissue=E-bryo, Larva, and Pupae;
MEDLINE=22426066; PubMed=12537569;
RA Stapleton X., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarnieri H., Krcmiller B., Pacle J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource."
Genome Biol. 3:RESEARCH080.1-RESEARCH080.8(2002).
CC Genome Biol. 3:RESEARCH080.1-RESEARCH080.8(2002).
CC -!- SIMILARITY: BELONGS TO THE SMALL HEAT SHOCK PROTEIN (HSP20)
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; J01098; AAF50290.1; -.
CC EMBL; AE003552; AAF50290.1; -.

DR EMBL; AY060412; AAL25451.1; -.
DR EMBL; AV119034; AAM50894.1; -.
DR PIR; A02918; HKFP22.
DR PIR; A20647; A20647.
DR FlyBase; FBgn0001223; Hsp22.
DR InterPro; IPR002068; Hsp20.
DR Pfam; PF0011; HSP20; 1.
DR PROSITE; PS01031; HSP20; 1.
KW Heat shock; Multigene family.
FT CONFLICT 44 44 H -> Q (IN REF. 3 AND 4).
FT CONFLICT 53 53 L -> F (IN REF. 3 AND 4).
FT CONFLICT 54 54 A -> P (IN REF. 2).
FT CONFLICT 90 90 A -> G (IN REF. 3 AND 4).
FT CONFLICT 109 111 RRF -> GRV (IN REF. 1).
FT CONFLICT 115 115 E -> D (IN REF. 1).
FT CONFLICT 123 128 TSTLSS -> SSSLSD (IN REF. 1).
FT CONFLICT 168 168 K -> N (IN REF. 3 AND 4).
FT CONFLICT 171 171 T -> A (IN REF. 3 AND 4).
SQ SEQUENCE 174 AA; 19796 MW; EEE10ACC37BE920 CRC64;

Query Match 50.6%; Score 41; DB 1; Length 174;
Best Local Similarity 54.5%; Pred. No. 20;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 HRPMPSPVMP 12
DB 29 HRPMPSPVMP 39

RESULT 5
VG27 BPML5 STANDARD; PRT; 336 AA.
AC Q05234;
DI 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DE 01-FEB-1994 (Rel. 28, Last annotation update)
DE Minor tail protein GP27.
GN 27.
OS Mycobacteriophage L5.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC L5-like viruses.
OX NCBI_TaxID=31757;
RN [1]
RP MEDLINE=93211282; PubMed=8459766;
RX Hatfull G.F., Sarkis G.J.;
RT "DNA sequence, structure and gene expression of mycobacteriophage L5:
RL Mol. Microbiol. 7:395-405(1993).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; Z18946; CAA79403.1; -.
DR PIR; S30972; S30972.
SQ SEQUENCE 336 AA; 38555 MW; 89EB47320D0809B8 CRC64;

Query Match 50.6%; Score 41; DB 1; Length 336;
Best Local Similarity 55.6%; Pred. No. 38;
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 PPMWSPVMP 12
DB 168 PPMWSPVMP 176

RESULT 6
NMT1_HUMAN


```

RN  [2]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=22388257; PubMed=2477932;
RA  Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA  Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA  Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA  Hopkins R.F., Jordan H., Moore T., Max S., Wang J., Hsieh F.,
RA  Diatchenko L., Matovina K., Farmer A.A., Rubin G.M., Hong L.,
RA  Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA  Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA  Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA  Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.R.,
RA  Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hu-Yk S.W.,
RA  Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA  Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA  Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA  Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA  Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA  Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA  Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT  "Generation and initial analysis of more than 15,000 full-length
RT  human and mouse cDNA sequences.";
RL  Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC  -!- FUNCTION: Adds a myristoyl group to the N-terminal glycine residue
CC  of certain cellular and viral proteins.
CC  -!- CATALYTIC ACTIVITY: Tetradecanoyl-CoA + glycyl-peptide = CoA + N-
CC  tetradecanoylglycyl-peptide
CC  -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC  -!- SIMILARITY: BELONGS TO THE NMT FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF043326; AAC09296.1; -
DR  EMBL; BC016526; AAH16526.1; -
DR  HSSP; P30418; 1NMT.
DR  MGSD; MGI:102579; Nmt1.
DR  InterPro; IPR000903; Nmt.
DR  Pfam; PF01233; NMT; 1.
DR  Pfam; PF02799; NMT; 1.
DR  PROSITE; PS00975; NMT_1; 1.
DR  PROSITE; PS00976; NMT_2; 1.
DR  TRANSFERASE; Acyltransferase.
KW  DOMAIN 55 67 POLY-LYS.
FT  CONFLICT 114 115 KR -> AK (IN REF. 2).
FT  CONFLICT 122 122 T -> V (IN REF. 2).
FT  CONFLICT 168 169 VL -> LF (IN REF. 2).
SQ  SEQUENCE 496 AA; 56888 MW; 61FA7B854A5CF3BC CRC64;
Query Match 50.6%; Score 41; DB 1; Length 496;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 RPPMWSPPV 11
Db 202 RPPGWLPPQW 210
RESULT 8
NMT1_BOVIN STANDARD; PRT; 497 AA.
AC P31717; Q9N177;
DT 01-JUL-1993 (Rel. 26, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Glycylpeptide N-tetradecanoyltransferase i (EC 2.3.1.97) (Peptide N-
DE myristoyltransferase 1) [Myristoyl-CoA:protein N-myristoyltransferase
DE 1] (NMT 1) (Type 1 N-myristoyltransferase).
GN NMT1 OR NMT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Retina;
RA Rundie D.R., Alvarez R.A., Anderson R.E.;
RT "Bovine retina type I N-myristoyltransferase.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 95-104; 114-123; 168-191 AND 449-459.
RX TISSUE=Brain;
RX MEDLINE=93199512; PubMed=9452528;
RA McIlhinney R.A.J., McGlone K., Willis A.C.;
RT "Purification and partial sequencing of myristoyl-CoA:protein N-
RT myristoyltransferase from bovine brain.";
RL Biochem. J. 290:405-416(1993).
CC -!- FUNCTION: Adds a myristoyl group to the N-terminal glycine residue
CC of certain cellular and viral proteins.
CC -!- CATALYTIC ACTIVITY: Tetradecanoyl-CoA + glycyl-peptide = CoA + N-
CC tetradecanoylglycyl-peptide.
CC -!- SUBUNIT: MAY BE ASSOCIATED WITH OTHER PROTEINS.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- PTM: THE N-TERMINUS IS BLOCKED.
CC -!- SIMILARITY: BELONGS TO THE NMT FAMILY.
CC  -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF223384; AAF31460.1; -
DR  HSSP; P30418; 1NMT.
DR  InterPro; IPR000903; Nmt.
DR  Pfam; PF01233; NMT; 1.
DR  Pfam; PF02799; NMT; 1.
DR  PROSITE; PS00975; NMT_1; 1.
DR  PROSITE; PS00976; NMT_2; 1.
DR  TRANSFERASE; Acyltransferase.
KW  DOMAIN 55 67 POLY-LYS.
FT  CONFLICT 114 115 KR -> AK (IN REF. 2).
FT  CONFLICT 122 122 T -> V (IN REF. 2).
FT  CONFLICT 168 169 VL -> LF (IN REF. 2).
SQ  SEQUENCE 497 AA; 56919 MW; 701A5C0B3554BC37 CRC64;
Query Match 50.6%; Score 41; DB 1; Length 497;
Best Local Similarity 66.7%; Pred. No. 55;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 3 RPPMWSPPV 11
Db 202 RPPGWLPPQW 210
RESULT 9
YH09_RALSO STANDARD; PRT; 657 AA.
AC Q8XV29;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein RSC1709.
GN RSC1709 OR RS02895.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN-GM11000;

```


CC between the Swiss Institute of Bioinformatics and the EMBL outstation:
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L07870; AAA35167.1; -
 DR EMBL: U22341; AAB60289.1; -
 DR EMBL: Z47815; CAA9781.1; -
 DR PIR: S50916; S50918.
 DR PDB: 1DB8; 13-JAN-83.
 DR SGD: S0004832; SGS1.
 DR GO: GO:0005730; C:nucleolus; IDA.
 DR GO: GO:0040003; P:ATP dependent DNA helicase activity; IDA.
 DR GO: GO:0007001; P:chromosome organization and biogenesis (sen. . .); IMP.
 DR GO: GO:0006268; P:DNA unwinding; IDA.
 DR GO: GO:0045132; P:mitotic chromosome segregation; IMP.
 DR GO: GO:0000070; P:mitotic chromosome segregation; IMP.
 DR InterPro: IPR001410; DEAD.
 DR InterPro: IPR002464; DEAH box.
 DR InterPro: IPR001650; Helicase_C.
 DR InterPro: IPR002121; HRDC.
 DR InterPro: IPR004589; RecQ.
 DR Pfam: PF00270; DEAH; 1.
 DR Pfam: PF00271; helicase_C; 1.
 DR Pfam: PF00570; HRDC; 1.
 DR SMART: SM00487; DEXDC; 1.
 DR SMART: SM00490; HELIC; 1.
 DR SMART: SM00342; HRDC; 1.
 DR TIGRfams: TIGR00614; recQ; 1.
 DR PROSITE: PS00690; DEAH ATP HELICASE; UNKNOWN 1.
 KW Helicase; ATP-binding; Nuclear protein; 3D-structure.
 FT DOMAIN 631 639 ASP/GLU-RICH (HIGHLY ACIDIC).
 FT NP-BIND 714 721 ATP (BY SIMILARITY).
 FT SITE 808 811 DEAH BOX.
 SQ SEQUENCE 1447 AA; 163836 MW; 0DC320B41756A3C3 CRC64;
 Query Match 50.6%; Score 41; DB 1; Length 1447;
 Best Local Similarity 75.8%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 2 HRPMPWSP 9
 DB 611 HRPMPWSP 618
 RESULT 12
 ID TTY7 HUMAN STANDARD; PRT; 124 AA.
 AC Q9BQ32;
 DT 15-SEP-2003 (Rel. 42, Created);
 DT 15-SEP-2003 (Rel. 42, Last sequence update);
 DE Transcription factor Y protein (fragment).
 GN TTY7 OR TTY7.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RA Kuroda-Kawaguchi T., Skaletsky H., Minx P.J., Brown L.G., Rozer S.,
 RA Wilson R.K., Waterston R.H., Page D.C.;
 RT "The DNA sequence of the human Y chromosome";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF332232; AAK13482.1; -
 DR EMBL: AF332233; AAK13483.1; -
 DR Genew; HGNC:18488; TTY7.
 FT NON TER 1
 SQ SEQUENCE 124 AA; 13790 MW; 4925F492546429B8 CRC64;
 Query Match 50.0%; Score 40.5; DB 1; Length 124;
 Best Local Similarity 53.8%; Pred. No. 17;
 Matches 7; Conservative 2; Mismatches 3; Indels 1; Gaps 1;
 QY 1 TERPPMW-SPWSP 12
 DB 108 TTKAPLWSPSWP 120
 RESULT 13
 ID DSB8 XANAC STANDARD; PRT; 172 AA.
 AC Q8PNQ6;
 DT 28-FEB-2003 (Rel. 41, Created);
 DT 28-FEB-2003 (Rel. 41, Last sequence update);
 DT 28-FEB-2003 (Rel. 41, Last annotation update);
 DE Disulfide bond formation protein B (Disulfide oxidoreductase).
 GN DSB8 OR XAC0998.
 OS Xanthomonas axonopodis (pv. citri).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=92829;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=306 / ATCC 13902 / XV 101;
 RX MEDLINE=2202145; PubMed=12024217;
 RA da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
 RA Canarotte G., Cannavan F., Cardoso J., Chambergo F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorfy H.,
 RA Fortighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
 RA Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities";
 RL Nature 417:459-463 (2002).
 CC -!- FUNCTION: Required for disulfide bond formation in some
 CC periplasmic proteins. Acts by oxidizing the dSbA protein (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (By similarity).
 CC -!- SIMILARITY: Belongs to the dSbB family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE011729; AAM3581.1; -
 DR HAMAP; MF_00286; -; 1.
 DR InterPro; IPR003752; DsbB.
 DR Pfam; PF02600; DsbB; 1.
 KW Oxidoreductase; Redox-active center; Transport; Electron transport;

KW Chaperone; Transmembrane; Inner membrane; Complete proteome.
 FT DOMAIN 1 1
 FT TRANSMEM 2 28
 FT DOMAIN 29 46
 FT TRANSMEM 47 63
 FT DOMAIN 64 70
 FT TRANSMEM 71 88
 FT DOMAIN 89 145
 FT TRANSMEM 146 164
 FT DOMAIN 165 172
 FT TRANSMEM 172 172
 FT DISULFID 38 41
 FT DISULFID 104 131
 SQ SEQUENCE 172 AA; 18990 MW; 503EA543B0F098A8 CRC64;

Query Match 49.4%; Score 40; DB 1; Length 172;
 Best Local Similarity 85.7%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PMWSPVW 11
 DB 144 PMWSMVW 150

RESULT 14
 DSBX_XANCP STANDARD; PRT; 172 AA.
 AC Q8PC32;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Disulfide bond formation protein B (Disulfide oxidoreductase).
 GN DSBX OR XCC0921.
 OS Xanthomonas campestris (pv. campestris).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 OC Xanthomonadaceae; Xanthomonas.
 OX NCBI_TaxID=340;
 RN 1;
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 33913 / NCPPB 528;
 RX MEDLINE=2202145; PubMed=12024217;
 RA da Silva A.C.R., Ferro C.A., Reinach F.C., Farah C.S., Furlan L.R.,
 RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
 RA Alves L.M.C., do Amaral A.M., Bertoloni M.C., Camargo L.E.A.,
 RA Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
 RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorri H.,
 RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferris M.Z.T.,
 RA Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
 RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
 RA Locati E.C., Machado M.A., Vadeira A.M.B.N., Martinez-Rossi N.M.,
 RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Noon D.H.,
 RA Moreira L.V., Novo T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
 RA Pereira H.A., Rossi A., Sena J.A.D., Silva C., de Souza R.F.,
 RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
 RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
 RA Setubal J.C., Kitajima J.P.;
 RT "Comparison of the genomes of two Xanthomonas pathogens with differing
 RT host specificities."
 RL Nature 417:459-463(2002).
 CC -!- FUNCTION: Required for disulfide bond formation in some
 CC periplasmic proteins. Acts by oxidizing the dsbA protein (by
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (by similarity).
 CC -!- SIMILARITY: Belongs to the dsbB family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR EMBL; AE012190; AA040231.1; -.
 DR HAMAP; MF 00286; -; 1.
 DR InterPro; IPR003752; Dsdb.
 DR Pfam; PF02600; Dsdb; 1.
 KW Oxidoreductase; Redox-active center; Transport; Electron transport;
 FT Chaperone; Transmembrane; Inner membrane; Complete proteome.
 FT DOMAIN 1 11
 FT TRANSMEM 12 28
 FT DOMAIN 29 46
 FT TRANSMEM 47 63
 FT DOMAIN 64 70
 FT TRANSMEM 71 88
 FT DOMAIN 89 145
 FT TRANSMEM 146 164
 FT DOMAIN 165 172
 FT DISULFID 38 41
 FT DISULFID 104 131
 SQ SEQUENCE 172 AA; 18931 MW; 541681DAEEF398E CRC64;

Query Match 49.4%; Score 40; DB 1; Length 172;
 Best Local Similarity 85.7%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PMWSPVW 11
 DB 144 PMWSMVW 150

RESULT 15
 UBCX_YEAST STANDARD; PRT; 183 AA.
 ID UBCX_YEAST
 AC P29340.
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Ubiquitin-conjugating enzyme E2-21 kDa (EC 6.3.2.19)
 DE (Ubiquitin-protein ligase) (Ubiquitin carrier protein) (Peroxin-4).
 GN PEX4 OR PAS2 OR UBC10 OR YGR133W.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN 1;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92396219; PubMed=1326082;
 RA Wiebel F.F., Kunau W.-H.,
 RA "The fast protein essential for peroxisome biogenesis is related to
 RA ubiquitin-conjugating enzymes."
 RL Nature 359:73-76(1992).
 RN 2;
 RP SEQUENCE FROM N.A.
 RA van Dyck L., Skala J., de Wergifosse P., Purnelle B., Talla E.,
 RA Nawrocki A., del Bino S., Goffeau A.,
 RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO
 CC OTHER PROTEINS. ESSENTIAL FOR PEROXISOME BIOGENESIS.
 CC -!- CATALYTIC ACTIVITY: ATP + ubiquitin + protein lysine = AMP +
 CC diphosphate + protein N-ubiquityllysine.
 CC -!- PATHWAY: Ubiquitin conjugation; second step.
 CC -!- SUBCELLULAR LOCATION: Peroxisomal.
 CC -!- MISCELLANEOUS: A cysteine residue is required for ubiquitin-
 CC thiolester formation.
 CC -!- SIMILARITY: Belongs to the ubiquitin-conjugating enzyme family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR EMBL; X65470; CAA46467.1; -.

DR EMBL; 272918; CAA9746.1; -.
 DR PIR; S29088; S29088.
 DR HSSP; P15731; IQCQ.
 DR SGD; SC003385; PEX4.
 DR InterPro; IPR000608; UBQ_conjugat.
 DR Pfam; PF03179; UQ_con; 1.
 DR ProDom; PDOC0461; UQ_conjugat; 1.
 DR SMART; SM0212; UBQC; 1.
 DR PROSITE; PS00183; UBIQUITIN_CONJUGAT_1; 1.
 DR PROSITE; PS0127; UBIQUITIN_CONJUGAT_2; 1.
 KW Ub1 conjugation pathway; Ligase; Multigene family; Peroxisome.
 FT BINDING 115 115 UBIQUITIN (BY SIMILARITY).
 SQ SEQUENCE 183 AA; 21119 MW; D4E438B689F76CAD CRC64;

Query Match 43.4%; Score 40; DB 1; Length 183;
 Best Local Similarity 55.6%; Pred.No. 29;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 RPFMWSFVW 11
 : : : : :
 Db 120 KPEEWTFVW 126

Search completed: October 22, 2003, 20:38:47
 Job time : 25 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 20:35:36 ; Search time 93 Seconds
(without alignments)
33.297 Million c.e.i. updates/sec

Title: US-09-995-804A-2

Perfect score: 81

Sequence: 1 THRPVWSPVMP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 3

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 23:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhs:*

8: sp_organelle:*

9: sp_phase:*

10: sp_plant:*

11: sp_rodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriaph:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	58.0	503	16	Q9HUL0 pseudomonas
2	46	56.8	162	16	Q9RVB6
3	45.5	56.2	546	5	O15919 trypanosoma
4	45.5	56.2	698	16	Q8ZJ11 yersinia pe
5	44.5	54.9	876	16	Q93IG6 mycobacteri
6	44	54.3	96	6	Q9BGS6 macaca fasc
7	44	54.3	124	11	Q9CGG5
8	44	54.3	414	11	Q8BYU9 mus musculu
9	44	54.3	475	11	Q8R268 mus musculu
10	44	54.3	476	11	Q8R268 mus musculu
11	44	54.3	946	10	Q81427 arabidopsi
12	44	54.3	1197	11	Q9Z0E7
13	44	54.3	1197	11	Q9R1X4 mus musculu
14	44	54.3	1472	11	Q8C0U5 mus musculu
15	43	53.1	233	17	Q58491 methanococ
16	43	53.1	503	3	Q8X0V8 neurospora

17	43	53.1	554	16	Q8ZD55
18	43	53.1	659	4	Q9NVK2
19	43	53.1	827	10	Q8M0A1
20	43	53.1	1183	2	Q8GCD1
21	43	53.1	1399	4	Q9P2D4
22	42.5	52.5	279	11	Q9CSJ8
23	42	51.9	116	5	Q9VQM2
24	42	51.9	383	10	Q8H024
25	42	51.9	454	10	Q9SDF4
26	42	51.9	485	4	Q8N8L1
27	42	51.9	720	2	Q8E089
28	42	51.9	791	10	Q8S982
29	42	51.9	826	10	Q8LQ00
30	42	51.9	3105	12	Q8U2I9
31	41.5	51.2	196	10	Q8GVM1
32	41.5	51.2	410	17	Q9HMJ7
33	41.5	51.2	437	16	Q9PFJ8
34	41.5	51.2	966	17	Q26770
35	41.5	51.2	1221	10	Q9FHR5
36	41	50.6	65	6	Q29588
37	41	50.6	104	10	Q9SHX3
38	41	50.6	122	10	Q8RUE2
39	41	50.6	184	10	Q9S784
40	41	50.6	221	5	Q61477
41	41	50.6	286	5	Q95XA8
42	41	50.6	300	2	Q9Z1I9
43	41	50.6	307	5	Q61478
44	41	50.6	329	16	Q53927
45	41	50.6	347	10	Q9ZRW7

ALIGNMENTS

RESULT 1

Q9HUL0 PRELIMINARY; PRT; 503 AA.

AC Q9HUL0;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)

DE Hypothetical protein PA4955.

GN PA4955.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;

OC Pseudomonadaceae; Pseudomonas.

OX NCBI_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 15692 / PA01;

RX MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,

RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuar Y.,

RA Brody L.A., Coulter S.N., Folger K.R., Kas A., Zarbig K., Lim R.M.,

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,

RA Reiter J., Sater M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

RT opportunistic pathogen.";

RJ Nature 406:959-964(2000).

DR EMBL; AE004908; AAG08340.1; ..

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 503 AA; 56255 MW; 430C0F83F1C4F40 CRC64;

Query Match 58.0%; Score 47; DB 16; Length 503;

Best Local Similarity 50.0%; Pred. No. 40;

Matches 7; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 3 RPPMWSP-----VWP 12

|||||

Db 303 RPPLWHPVALIWP 316

RESULT 2

Q9RVB6 PRELIMINARY; PRT; 162 AA.
 AC Q9RVB6;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein DR113.
 GN DR113.
 OS Deinococcus radiodurans.
 OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=R1;
 RA MEDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Voffat K.S., Qin H., Jiang L., Bamphile W., Crosby M., Shen M.,
 RA Vanathevan J.J., Lam P., McDonald L., Uterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RA "Genome sequence of the radioresistant bacterium Deinococcus
 RT radiodurans R1";
 RL Science 286:1571-1577(1999).
 DR EMBL; AE000961; AAF0691.1; -.
 DR TIGR; DR113; -.
 DR InterPro; IPR006750; DUF606.
 DR Pfam; PF04657; DUF606; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 162 AA; 16547 MW; 59FC43C0C8C43557B CRC64;

Query Match: 56.2%; Score 46; DB 16; Length 162;
 Best Local Similarity 60.0%; Pred. No. 20;

Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 HRPMPWSPVW 11
 ||| |
 DB 63 HRPDHPWPLW 72

RESULT 3

O15919 PRELIMINARY; PRT; 546 AA.
 AC O15919;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Mitochondrial DEAD box protein.
 GN VHEL61.
 OS Trypanosoma brucei.
 OC Eukaryota; Eucenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5691;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1Star 1;
 RA MEDLINE=97413570; PubMed=92711369;
 RA Misset A., Souza A.E., NorSkau G., Goring H.U.;
 RA "Disruption of a gene encoding a novel mitochondrial DEAD-box protein
 RT in Trypanosoma brucei affects edited mRNAs";
 RJ Mol. Cell. Biol. 17:4895-4903(1997).
 CC -1- SIMILARITY: BELONGS TO THE DEAD BOX HELICASE FAMILY.
 DR EMBL; U86382; AAB69639.1; -.
 DR HSSP; Q58083; 1HV8.
 DR InterPro; IPR001410; DEAD.
 DR InterPro; IPR000629; DEAD box.
 DR Pfam; PF00270; DEAD; 1.
 DR Pfam; PF00271; helicase_C; 1.
 DR SMART; SM00487; DEXDC; 1.
 DR SMART; SM00490; HELICC; 1.

DR PROSITE; PS00339; DEAD_ATP_HELICASE; 1.
 KW ATP-binding; Helicase; Hydrolase.
 SQ SEQUENCE 546 AA; 60719 MW; 228E6E008910F63A CRC64;

Query Match: 56.2%; Score 45.5; DB 5; Length 546;
 Best Local Similarity 53.3%; Pred. No. 71;
 Matches 8; Conservative 1; Mismatches 3; Indels 3; Gaps 1;

QY 1 THRPV---MWSVPWP 12
 ||| |
 DB 290 SHRPACTMWSATWP 304

RESULT 4

Q8ZJ11 PRELIMINARY; PRT; 698 AA.
 AC Q8ZJ11;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE 4-alpha-glucanotransferase (EC 2.4.1.25) (Amylomaltase).
 GN MALQ OR YP00126 OR Y3902.
 OS Versinia pestis.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Versinia.
 OX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / Biovar Orientalis;
 RX MEDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Taitball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Fellwell T., Hamlin N., Ho-royd S., Jagels K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrell B.G.;
 RC "Genome sequence of Versinia pestis, the causative agent of plague";
 RL Nature 413:523-527(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KIM5 / Biovar Mediaevalis;
 RX MEDLINE=22137863; PubMed=12142430;
 RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
 RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
 RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
 RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
 RA Perry R.D.;
 RT "Genome sequence of Versinia pestis KIM";
 RL J. Bacteriol. 184:4601-4611(2002).
 DR EMBL; AJ41414; CAC88989.1; -.
 DR EMBL; AE013995; AAX87446.1; -.
 DR InterPro; IPR003385; Glyco_hydro_77.
 DR Pfam; PF02446; 4A_glucanotrans; 1.
 DR TIGRFAMS; TIGR00217; malQ; 1.
 KW Glycosyltransferase; Transferase; Complete proteome.
 SQ SEQUENCE 698 AA; 78553 MW; 685512C27B2A692 CRC64;

Query Match: 56.2%; Score 45.5; DB 16; Length 698;
 Best Local Similarity 66.7%; Pred. No. 89;

Matches 8; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 2 HRPVWWS-VVWP 12
 ||| |
 DB 313 HDPMMWGVVWP 324

RESULT 5

Q93IG6 PRELIMINARY; PRT; 876 AA.
 ID Q93IG6
 AC Q93IG6;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein RV3365c.
 GN RV3365C OR MT3474.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigemeier K., Gas S., Barry III C.E., Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Harlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter K., Seeger K., Skelton S., Squares S., Sgares R., Suiston J.E.,
 RA Taylor K., Whitehead S., Barrrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RA Parkhill J.;
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.B., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva X.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Nikula A.,
 RA Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL003198; CRA15750.1; -;
 DR ENBL: AE007154; AAK47812.1; -;
 DR TIGR: MT3474; -;
 DR TubercuList; RV3365c; -;
 DR InterPro; IPR003594; ATPbind_ATPase.
 DR InterPro; IPR003660; HAMF.
 DR InterPro; IPR006025; 2r_MTPpeptidase.
 DR Pfam; PF00672; HAMF; 1.
 DR Pfam; PF02518; HAMFase_C; 1.
 DR SMART; SM00304; HAMF; 1.
 DR PROSITE; PS00142; ZINC_PROTEASE; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 876 AA; 93441 MW; D53EDAFDCD12CA5 CRC64;
 Query Match 54.9%; Score 44.5; DB 16; Length 876;
 Best Local Similarity 72.7%; Pred. No. 1.5e+02;
 Matches 8; Conservative 0; Mismatches 2; Indels 1; Gaps 1;
 QY 3 RPPWSPPVWP 12
 DB 32 RPPWSPPVWP 42
 RESULT 6
 Q9BGS6 PRELIMINARY; PRT; 96 AA.
 ID Q9BGS6
 AC Q9BGS6;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Hypothetical 10.2 kDa protein.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;

OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9541;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Frontal cortex;
 RA Oseada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain: cDNA
 RT libraries."
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB056398; BAB33054.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 96 AA; 10190 MW; C95C6C326952A5B2 CRC64;
 Query Match 54.3%; Score 44; DB 6; Length 96;
 Best Local Similarity 60.0%; Pred. No. 23;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
 QY 3 RPPWSPVWP 12
 DB 15 RPPWSPVWP 24
 RESULT 7
 Q9CQG5 PRELIMINARY; PRT; 124 AA.
 ID Q9CQG5
 AC Q9CQG5;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE 473340:F03Rik protein.
 GN 473340:F03Rik.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Skin, Embryonic stem cells, and Embryo;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Guscinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690 (2001).
 DR EMBL: AK014632; BAB29475.1; -;
 DR EMBL: AK010798; BAB27188.1; -;
 DR EMBL: AK013126; BAB28664.1; -;
 DR MGD; MGI:1913960; 473340:F03Rik.
 SQ SEQUENCE 124 AA; 14519 MW; 9041EA9268D438FB CRC64;
 Query Match 54.3%; Score 44; DB 11; Length 124;
 Best Local Similarity 58.3%; Pred. No. 30;
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
 QY 1 THRPWSPVWP 12
 DB 87 TYSPPWSPVWP 98


```

RESULT 8
QB8YU9
ID QB8YU9 PRELIMINARY; PRT; 414 AA.
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE SECTs-binding protein 2 homolog.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL: AK037988; BAC29914.1; -.
SQ SEQUENCE 414 AA; 45156 MW; F9F1B46A62A8B2B3 CRC64;

Query Match 54.3%; Score 44; DB 11; Length 414;
Best Local Similarity 50.0%; Pred. No. 90;
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 THRPWMSPWP 12
Db 238 TORPPRWGLGP 249

RESULT 9
QB8YU9
ID QB8YU9 PRELIMINARY; PRT; 475 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Timeless protein small isoform.
GN TIMELESS.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=20420328; PubMed=10963667;
RA Li Z., Stuart R.O., Qiao J., Pavlova A., Bush K.T., Pohl M.,
RA Sakurai H., Nigam S.K.;
RT "A role for Timeless in epithelial morphogenesis during kidney
RT development."
RL Proc. Natl. Acad. Sci. U.S.A. 97:10038-10043(2000).
DR EMBL: AF126480; AJ224467.1; -.
RX MGD; MGI:1321393; Timeless.
DR InterPro; IPR001005; Myb_DNA_binding.
DR PROSITE; PS00037; MYB_1.1.
SQ SEQUENCE 475 AA; 53980 MW; 930C1D8832B0802A CRC64;

Query Match 54.3%; Score 44; DB 11; Length 475;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 THRPWMSPWP 9
Db 89 SHRAPLWSP 97

RESULT 10
QB8YU9
ID QB8YU9 PRELIMINARY; PRT; 476 AA.
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Similar to timeless homolog (Drosophila).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC026526; AAH26526.1; -.
DR InterPro; IPR001005; Myb_DNA_binding.
DR PROSITE; PS00037; MYB_1.1.
SQ SEQUENCE 476 AA; 54014 MW; 3B6BA2FB91C9ECE0 CRC64;

Query Match 54.3%; Score 44; DB 11; Length 476;
Best Local Similarity 66.7%; Pred. No. 1e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 THRPWMSPWP 9
Db 89 SHRAPLWSP 97

RESULT 11
QB8YU9
ID QB8YU9 PRELIMINARY; PRT; 946 AA.
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE T24H24.6 protein (Putative transposon protein).
GN T24H24.6 OR A14G04140.
CS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA WASHU;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Courtney L., Stoneking T., Langston Y., Mead K.;
RT "The sequence of A. thaliana T24H24."
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Waterston R.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Larar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF075598; AAC28204.1; -.
DR EMBL; AL161499; CAB77882.1; -.
DR InterPro; IPR004332; MUDR.

```

DR InterPro: IPR000531; TonB_boxC.
 DR InterPro: IPR006564; Znf_PNZ.
 DR Pfam: PF03108; ZNF, 1.
 DR SMART: SMC0575; ZNF_PNZ, 1.
 DR PROSITE: PS00430; TONB_DEPENDENT_REC_1; 1.
 SQ SEQUENCE 946 AA; 106331 MW; 20E9967299F2B194 CRC64;

Query Match 54.3%; Score 44; DB 11; Length 946;
 Best Local Similarity 87.5%; Pred. No. 1.9e+02;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 HRPMPWSP 9
 ||| |||
 Db 746 HRLPWSP 753

RESULT 12
 Q920E7 PRELIMINARY; PRT; 1:97 AA.
 ID Q920E7
 AC Q920E7
 DT 01-MAY-1999 (TRENBLrel. 10, Created;
 DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE MTM1.
 GN TIMELESS OR MTM1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BAJB/C;
 RX MEDLINE=93107002; PubMed=9891984;
 RA Koike N., Hida A., Numano R., Hirose M., Sakaki Y., Tei K.;
 RT "Identification of the mammalian homologues of the Drosophila timeless
 RT gene, Timeless1.";
 RL FEBS Lett. 441:427-431(1998).
 RN [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=99072286; PubMed=9856466;
 RA Zylka M.J., Shearman L.P., Levine J.D., Jin X., Weaver D.R.,
 RA Reppert S.M.;
 RT "Molecular analysis of mammalian timeless.";
 RL Neuron 21:1115-1122(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99072285; PubMed=9856465;
 RA Sangoram A.M., Saez L., Antoch M.P., Gekakis N., Staknis D.,
 RA Whiteley A., Fruechte E.M., Vitaterna M.H., Shimomura K., King D.P.,
 RA Young M.W., Weitz C.J., Takahashi J.S.;
 RT "Mammalian circadian autoregulatory loop: a timeless ortholog and
 RT mPer1 interact and negatively regulate CLOCK-BMAL1-induced
 RT transcription.";
 RL Neuron 21:1101-1113(1998).
 DR EMBL; AF015598; BAA36500.1; -;
 DR EMBL; AF015506; AAC79687.1; -;
 DR EMBL; AF098161; AAC80010.1; -;
 DR MGD; MGI:1321393; Timeless.
 DR InterPro: IPR001005; Myb_DNA_binding.
 DR InterPro: IPR006906; Timeless.
 DR Pfam; PF04821; TIMELESS; 1.
 DR PROSITE; PSC0037; MYB_1; 1.
 SQ SEQUENCE 1197 AA; 137502 MW; DB24B7F0F3D05787 CRC64;

Query Match 54.3%; Score 44; DB 11; Length 1197;
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 THRPMPWSP 9
 ||| |||
 Db 810 SHRAPLWSP 818

RESULT 13
 Q9R1X4 PRELIMINARY; PRT; 1:97 AA.
 ID Q9R1X4
 AC Q9R1X4
 DT 01-MAY-2000 (TRENBLrel. 13, Created)
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Timeless protein (Fragment).
 GN TIMELESS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99248428; PubMed=10231394;
 RA Takumi T., Nagamine Y., Miyake S., Matsubara C., Taguchi K.,
 RA Takekida S., Sakakida Y., Nishikawa K., Niwa S., Okumura H.;
 RT "A mammalian ortholog of drosophila timeless, highly expressed in SCN
 RT and retina, forms a complex with mPER1.";
 RL Genes Cells 4:67-75(1999).
 DR EMBL; AB019001; BAA76390.2; -;
 DR MGD; MGI:1321393; Timeless.
 DR InterPro: IPR001005; Myb_DNA_binding.
 DR InterPro: IPR006906; Timeless.
 DR Pfam; PF04821; TIMELESS; 1.
 DR PROSITE; PSC0037; MYB_1; 1.
 FT NON_TER 1197 1197
 SQ SEQUENCE 1197 AA; 137438 MW; 67725B03B2C0F4D CRC64;

Query Match 54.3%; Score 44; DB 11; Length 1197;
 Best Local Similarity 66.7%; Pred. No. 2.4e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 THRPMPWSP 9
 ||| |||
 Db 810 SHRAPLWSP 818

RESULT 14
 Q8C0U5 PRELIMINARY; PRT; 1472 AA.
 ID Q8C0U5
 AC Q8C0U5
 DT 01-MAR-2003 (TRENBLrel. 23, Created)
 DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
 DE Hypothetical dynein heavy chain containing protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Testis;
 RX MEDLINE=22354683; PubMed=12466851;
 RA The PANTOM Consortium.
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 DR EMBL; AK029791; BAC26619.1; -;
 KW Hypothetical protein.
 SQ SEQUENCE 1472 AA; 168538 MW; EDF5B035DE2AA96D CRC64;

Query Match 54.3%; Score 44; DB 11; Length 1472;
 Best Local Similarity 54.5%; Pred. No. 2.9e+02;
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 HRPMPWSPVWP 12
 ||| |||
 Db 1280 HVPPLWGRVYP 1290

```

RESULT 15
Q58491 PRELIMINARY; PRT; 233 AA.
AC Q58491;
DT 01-JAN-1998 (TRENBLRE: 05, Created)
DT 01-JAN-1998 (TRENBLRE: 05, Last sequence update)
DT 01-JUN-2002 (TRENBLRE: 21, Last annotation update)
DE Cobalt transport ATP-binding protein CblM.
GN CblM OR Mj1091.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=868087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hust M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RL Science 273:1058-1073(1996).
CC -!- FUNCTION: NOT KNOWN.
CC -!- PATHWAY: COBALAMIN BIOSYNTHESIS.
DR EMBL; U67551; AAB99092.; -.
DR TIGR; Mj1091; -.
DR InterPro; IPR002751; CblM.
DR Pfam; PF01892; CblM; 1.
DR ProDom; PD005331; CblM; 1.
DR TrGFAMS; TIGR00123; cblM; 1.
KW Cobalamin biosynthesis; Transmembrane; Complete proteome.
FT TRANSMEM 9 29 POTENTIAL.
FT TRANSMEM 43 63 POTENTIAL.
FT TRANSMEM 75 95 POTENTIAL.
FT TRANSMEM 107 27 POTENTIAL.
FT TRANSMEM 138 158 POTENTIAL.
FT TRANSMEM 177 197 POTENTIAL.
SQ SEQUENCE 233 AA; 24914 MW; 7DC194B58AA2CF93 CRC64;

Query Match 53.1%; Score 43; DS 17; Length 233;
Best Local Similarity 75.0%; Pred. No. 74;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PPMWSPVW 11
Db 9 PPMWCAVW 16

Search completed: October 22, 2003, 20:40:29
Job time : 97 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 20:41:21 ; Search time 68 seconds
(without alignments)
29.552 Million cell updates/sec

Title: US-09-995-804A-2

Perfect score: 81

Sequence: 1 THRPWMSPVWP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 629382

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US05_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
17: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	12	10	US-09-995-804A-2
2	50	61.7	127	9	US-09-864-761-38565
3	47	58.0	127	15	US-10-106-698-8065
4	45.5	56.2	99	12	US-10-029-386-31023
5	44	54.3	20	12	US-09-962-756-1066
6	44	54.3	20	12	US-09-962-756-1255
7	44	54.3	20	12	US-09-962-756-1255
8	44	54.3	208	12	US-10-029-386-33817
9	44	54.3	1197	15	US-10-245-175-2
10	43	53.1	59	11	US-09-764-891-4960
11	42	51.9	59	9	US-09-864-761-37745
12	42	51.3	137	15	US-10-156-761-10941
13	42	51.9	554	15	US-10-184-128-2
14	42	51.9	963	12	US-10-073-333A-2
15	42	51.9	963	15	US-10-140-164-2

16	42	51.9	1027	12	US-10-073-333A-4
17	42	51.9	1027	15	US-10-140-164-4
18	41	50.6	89	9	US-09-925-297-646
19	41	50.6	178	15	US-10-106-698-4480
20	41	50.6	388	10	US-09-361-630-1
21	41	50.6	478	12	US-10-427-631-31
22	40	49.4	20	9	US-09-879-957-181
23	40	49.4	20	9	US-09-879-957-223
24	40	49.4	24	9	US-09-879-957-41
25	40	49.4	36	15	US-10-050-704-319
26	40	49.4	85	12	US-10-029-386-31584
27	40	49.4	88	9	US-09-764-860-386
28	40	49.4	88	15	US-10-074-095-386
29	40	49.4	135	15	US-10-050-704-325
30	40	49.4	177	15	US-10-050-704-316
31	40	49.4	305	9	US-09-925-302-620
32	40	49.4	429	12	US-10-029-386-33356
33	40	49.4	433	8	US-08-979-847-81
34	40	49.4	433	12	US-10-430-442-87
35	40	49.4	437	11	US-09-374-046A-66
36	40	49.4	578	15	US-10-156-761-7579
37	40	49.4	2364	15	US-10-156-761-7834
38	39.5	48.8	117	11	US-09-764-891-2956
39	39.5	48.8	142	15	US-10-125-258-52
40	39.5	48.8	142	15	US-10-125-258-53
41	39	48.1	33	9	US-09-864-761-38235
42	39	48.1	38	12	US-10-029-386-32715
43	39	48.1	55	15	US-10-267-849-43
44	39	48.1	65	9	US-09-864-761-36469
45	39	48.1	70	9	US-09-764-887-272

ALIGNMENTS

RESULT 1

US-09-804A-2
; Sequence 2, Application US/0995804A
; Patent No. US20020115824A1
; GENERAL INFORMATION:
; APPLICANT: Engler, Jeffrey A
; Lee, Jae Hwy
; Collawan, James F
; Moore, Bryan A

Self ref

TITLE OF INVENTION: Receptor-Mediated Uptake of Peptides
that Bind the Human Transferrin Receptor

NUMBER OF SEQUENCES: 13

CORRESPONDENCE ADDRESS:

ADDRESSEE: Hendricks and Assoc.

STREET: P.O. Box 2509

CITY: Fairfax

STATE: VA

COUNTRY: U.S.A

ZIP: 22031

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/995-804A

FILING DATE: 29-No. US20020115824A1-2001

CLASSIFICATION: Unknown

ATTORNEY/AGENT INFORMATION:

NAME: Hendricks, Glenn M

REGISTRATION NUMBER: 32,535

REFERENCE/DOCKET NUMBER: engler1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 703/425-8405

TELEFAX: 703/425-8406

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 12 amino acids

```
;
;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-995-804a-2

Query Match      100.0%; Score 81; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 THRPMPKSPVWP 12
       |||:|||||
Db      1 THRPMPMSPVWP 12

RESULT 2
US-09-864-761-38565
; Sequence 38565, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wesheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE OF INVENTION: Aescmca-X-1
; FILE REFERENCE: Aescmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2003-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2003-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,409
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 38565
; LENGTH: 127
; TYPE: PRT
```

```
;
;
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005962.1
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 6.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 19
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.2
; OTHER INFORMATION: SWISSPROT HIT: P21139, EVALU6 6.30e+00
US-09-864-761-38565

Query Match      61.7%; Score 50; DB 9; Length 127;
Best Local Similarity 54.5%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY      1 THRPMPMSPVW 11
       |||:|
Db      30 TRSPFLWTPAW 40

RESULT 3
US-10-106-698-8065
; Sequence 8065, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; FILE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypept
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,260
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 8065
; LENGTH: 127
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (105)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (108)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (111)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (113)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (126)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-8065

Query Match      58.0%; Score 47; DB 15; Length 127;
Best Local Similarity 63.6%; Pred. No. 39;
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      2 HRPPMSPVWP 12
       |||:|
Db      114 HEPPAWDPVAP 124

RESULT 4
```

US-10-029-386-31023
; Sequence 31023, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE OF INVENTION: EXPRESSION ANALYSIS TWO
; FILE REFERENCE: AECOMICA-X-2 US/10/029,386
; CURRENT APPLICATION NUMBER: US/10/029,386
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 31023
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL121906.15
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.96
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.92
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.3
US-10-029-386-31023
Query Match 56.2%; Score 45.5; DB 12; Length 99;
Best Local Similarity 70.0%; Pred. No. 50;
Matches 7; Conservative 1; Mismatches 1; Indels 1; Gaps 1;
QY 4 PPMWS-PVWP 12
DB 35 PPLWSEPAWP 44
US-09-962-756-1086
; Sequence 1086, Application US/09952756
; Publication No. US20030195147A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA
; APPLICANT: BRISSETTE, RENEE
; APPLICANT: BLUME, ARTHUR J.
; APPLICANT: SCHAFFER, LAUGE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4051J81
; CURRENT APPLICATION NUMBER: US/09/962,756
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,039
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1086
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-962-756-1086
Query Match 54.3%; Score 44; DB 12; Length 20;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 RPPMWSPVWP 12
DB 2 RVALWGPVWP 11
RESULT 6
US-09-962-756-1255
; Sequence 1255, Application US/09962756
; Publication No. US20030195147A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA
; APPLICANT: BRISSETTE, RENEE
; APPLICANT: BLUME, ARTHUR J.
; APPLICANT: SCHAFFER, LAUGE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4051J81
; CURRENT APPLICATION NUMBER: US/09/962,756
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1255
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: Peptide
US-09-962-756-1255
Query Match 54.3%; Score 44; DB 12; Length 20;
Best Local Similarity 60.0%; Pred. No. 21;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 3 RPPMWSPVWP 12
DB 2 RVALWGPVWP 11
RESULT 7
US-09-864-761-42563
; Sequence 42563, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: AECOMICA-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666

```
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42563
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ACC20904.3
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 4
; OTHER INFORMATION: ES: HUMAN HIT: AA62190.1, EVALUE 2.00e-24
; OTHER INFORMATION: SW:SSPROT HIT: F16759, EVALUE 3.60e+00
; OTHER INFORMATION: SW:SSPROT HIT: F16759, EVALUE 3.60e+00
US-09-864-761-42563

Query Match 54.3%; Score 44; DB 9; Length 51;
Best Local Similarity 85.7%; Pred. No. 45;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 PWSPPW 11
DB 4 PWSPPW 10

RESULT 8
US-10-029-386-33817
; Sequence 33817, Application US/13029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David P.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEWICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33817
; LENGTH: 208
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011547.3

; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
; OTHER INFORMATION: SW:SSPROT HIT: P15276, EVALUE 3.09e-09
US-10-029-386-33817

Query Match 54.3%; Score 44; DB 12; Length 208;
Best Local Similarity 60.0%; Pred. No. 1.4e+02;
Matches 6; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMSPW 12
DB 5 RPLWASVFP 14

RESULT 9
US-10-245-175-2
; Sequence 2, Application US/10245175
; Publication No. US20030059848A1
; GENERAL INFORMATION:
; APPLICANT: Reppert, Steven M.
; APPLICANT: Weaver, David R.
; APPLICANT: Zylka, Mark
; APPLICANT: Jin, Xiaowei
; APPLICANT: Kume, Kazuhiko
; APPLICANT: Sirram, Sathyanarayanan
; APPLICANT: Shearman, Lauren
; TITLE OF INVENTION: METHODS FOR IDENTIFYING COMPOUNDS WHICH
; TITLE OF INVENTION: MODULATE CIRCADIAN RHYTHM
; FILE REFERENCE: 00786-428002
; CURRENT APPLICATION NUMBER: US/10/245,175
; CURRENT FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: US 09/618,425
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: US 60/203,005
; PRIOR FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: US 60/145,363
; PRIOR FILING DATE: 1999-07-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1197
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-245-175-2

Query Match 54.3%; Score 44; DB 15; Length 1197;
Best Local Similarity 66.7%; Pred. No. 5.9e+02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 THRPMPSP 9
DB 810 SHRAPLWSP 819

RESULT 10
US-09-764-891-4960
; Sequence 4960, Application US/09764891
; Publication No. US20030077808A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4960
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
```

/ LOCATION: (20)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (27)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (53)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (56)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (58)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (59)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
/ NAME/KEY: SITE
/ LOCATION: (64)
/ OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-4960

Query Match 53.1%; Score 43; DB 11; Length 68;
Best Local Similarity 50.0%; Pred. No. 76;
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 3 RPPMSPVMP 12
||| |||
DB 14 KPPLWGXWMP 23

RESULT 11
US-09-864-761-37745
/ Sequence 37745, Application US/0986476;
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: Aeonica-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 63/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30

/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 37745
/ LENGTH: 59
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO AC002094.1
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 31
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 4.9
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 31
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 7.6
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 36
/ OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 11
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 8.8
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
/ OTHER INFORMATION: EST HUMAN HIT: AAC75986.1, EVALUE 2.00e-07
/ OTHER INFORMATION: SWISSPROT HIT: P05783, EVALUE 1.20e-02
US-09-864-761-37745

Query Match 51.9%; Score 42; DB 9; Length 59;
Best Local Similarity 75.0%; Pred. No. 92;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 PMSPVMP 12
||| |||
DB 20 PSWSAVMP 27

RESULT 12
US-10-156-761-10941
/ Sequence 10941, Application US/10156761
/ Publication No. US20030119018A1
/ GENERAL INFORMATION:
/ APPLICANT: OMURA, SATOSHI
/ APPLICANT: IKEDA, HARUO
/ APPLICANT: ISHIKAWA, JUN
/ APPLICANT: HORIKAWA, HIROSHI
/ APPLICANT: SHIBA, TADAYOSHI
/ APPLICANT: SAKAKI, YOSHIYUKI
/ APPLICANT: HATTORI, MASAHIRA
/ TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
/ FILE REFERENCE: 249-262
/ CURRENT APPLICATION NUMBER: US/10/156,761
/ CURRENT FILING DATE: 2002-05-29
/ PRIOR APPLICATION NUMBER: JP 2001-204089
/ PRIOR FILING DATE: 2001-05-30
/ PRIOR APPLICATION NUMBER: JP 2001-272697
/ PRIOR FILING DATE: 2001-08-02
/ NUMBER OF SEQ ID NOS: 15109
/ SEQ ID NO 10941
/ LENGTH: 137
/ TYPE: PRT
/ ORGANISM: Streptomyces avermitilis
US-10-156-761-10941

Query Match 51.9%; Score 42; DB 15; Length 137;
Best Local Similarity 69.2%; Pred. No. 1.8e+02;
Matches 9; Conservative 1; Mismatches 1; Indels 2; Gaps 2;
QY 1 THRPMPW-SP-VW 11
||| |||
DB 66 TRRPPLWASPRVW 78


```

RESULT 13
US-10-184-128-2
; Sequence 2, Application US/10184.128
; Publication No. US2003004971A1
; GENERAL INFORMATION:
; APPLICANT: Ebens Jr., Allen James
; APPLICANT: Stout, Thomas J.
; APPLICANT: Nguyen, Dat
; APPLICANT: Johnston, Stuart
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING INSECT CYSTEINE
; FILE REFERENCE: G03-009
; CURRENT APPLICATION NUMBER: US/10/184.128
; CURRENT FILING DATE: 2002-06-26
; PRIOR APPLICATION NUMBER: 60/301,968
; PRIOR FILING DATE: 2001-06-29
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Heliothis virescens
US-10-184-128-2

Query Match 51.9%; Score 42; DB 15; Length 554;
Best Local Similarity 75.0%; Pred. No. 5.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RPPMSPVW 11
DB 23 PPQWSPVY 30

RESULT 14
US-10-073-333A-2
; Sequence 2, Application US/10073333A
; Publication No. US20030134788A1
; GENERAL INFORMATION:
; APPLICANT: Kevin Baker et al.
; TITLE OF INVENTION: HUMAN TUMOR NECROSIS FACTOR RECEPTOR TR16
; FILE REFERENCE: PFS14PI
; CURRENT APPLICATION NUMBER: US/10/073.333A
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: 60/268,364
; PRIOR FILING DATE: 2001-02-24
; PRIOR APPLICATION NUMBER: 09/637,856
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 60/148,348
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/148,683
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/149,181
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149,453
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/149,498
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 963
; TYPE: PRT
; ORGANISM: human
US-10-073-333A-2

Query Match 51.9%; Score 42; DB 12; Length 963;
Best Local Similarity 66.7%; Pred. No. 8.9e+02;
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```

```

Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 RPPMSPVW 11
DB 25 RSPWSPAW 33

RESULT 15
US-10-140-164-2
; Sequence 2, Application US/10140164
; Publication No. US20030072736A1
; GENERAL INFORMATION:
; APPLICANT: Baker et al.
; TITLE OF INVENTION: Human Tumor Necrosis Factor Receptor TR16
; FILE REFERENCE: PFS14CI
; CURRENT APPLICATION NUMBER: US/10/140.164
; CURRENT FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: 09/637,856
; PRIOR FILING DATE: 2000-08-10
; PRIOR APPLICATION NUMBER: 60/148,348
; PRIOR FILING DATE: 1999-08-12
; PRIOR APPLICATION NUMBER: 60/148,683
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 60/148,870
; PRIOR FILING DATE: 1999-08-13
; PRIOR APPLICATION NUMBER: 60/148,758
; PRIOR FILING DATE: 1999-08-16
; PRIOR APPLICATION NUMBER: 60/149,181
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149,453
; PRIOR FILING DATE: 1999-08-18
; PRIOR APPLICATION NUMBER: 60/149,498
; PRIOR FILING DATE: 1999-08-19
; NUMBER OF SEQ ID NOS: 76
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 963
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-140-164-2

Query Match 51.9%; Score 42; DB 15; Length 963;
Best Local Similarity 66.7%; Pred. No. 8.9e+02;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 RPPMSPVW 11
DB 25 RSPWSPAW 33

Search completed: October 22, 2003, 20:50:30
Job time : 71 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Corpuen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 20:38:26 ; Search time 28 Seconds
(without alignments)
18.133 Million cell updates/sec

Title: US-09-995-804A-2

Sequence: 1 THRPPMWSVPW 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: /cgn2_6/ptodata/2/iaa/5A.COMB.pep.*
- 2: /cgn2_6/ptodata/2/iaa/5B.COMB.pep.*
- 3: /cgn2_6/ptodata/2/iaa/6A.COMB.pep.*
- 4: /cgn2_6/ptodata/2/iaa/6B.COMB.pep.*
- 5: /cgn2_6/ptodata/2/iaa/PCUTUS.COMB.pep.*
- 6: /cgn2_6/ptodata/2/iaa/backfiles.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48.5	59.9	603	4	US-09-252-991A-32674
2	47	58.0	521	4	US-09-252-991A-32732
3	45	55.6	145	4	US-09-252-991A-31281
4	44	54.3	1197	4	US-09-618-425-2
5	43	53.1	501	2	US-08-660-963-13
6	42.5	52.5	110	4	US-09-252-991A-20271
7	42	51.9	345	4	US-09-252-991A-25448
8	42	51.9	441	4	US-09-252-991A-28965
9	42	51.9	720	1	US-07-731-157A-2
10	42	51.9	720	2	US-08-541-780-2
11	41.5	51.2	469	4	US-09-252-991A-30596
12	41.5	51.2	882	4	US-09-252-991A-24619
13	41	50.6	160	4	US-09-252-991A-22258
14	41	50.6	266	4	US-09-252-991A-18776
15	41	50.6	478	4	US-09-786-240-3
16	41	50.6	2860	4	US-08-826-267-2
17	40.5	50.0	143	4	US-09-252-991A-17703
18	40.5	50.0	589	4	US-09-252-991A-18210
19	40	49.4	20	4	US-08-630-915A-181
20	40	49.4	22	4	US-08-630-915A-223
21	40	49.4	41	24	US-08-630-915A-41
22	40	49.4	30	2	US-08-753-829A-5
23	40	49.4	109	4	US-09-252-991A-29444
24	40	49.4	133	4	US-09-252-991A-23595
25	40	49.4	134	4	US-09-252-991A-22958
26	40	49.4	142	4	US-09-252-991A-17426
27	40	49.4	189	4	US-09-252-991A-16638

28	40	49.4	259	4	US-09-252-991A-28679	Sequence 28679, A
29	40	49.4	318	2	US-08-960-022-4	Sequence 4, Appli
30	40	49.4	318	3	US-09-058-368-3	Sequence 3, Appli
31	40	49.4	351	4	US-09-252-991A-21647	Sequence 21647, A
32	40	49.4	361	4	US-09-252-991A-18745	Sequence 18745, A
33	40	49.4	375	4	US-09-252-991A-24278	Sequence 24278, A
34	40	49.4	433	3	US-08-691-563C-87	Sequence 87, Appli
35	40	49.4	433	4	US-09-374-766-87	Sequence 87, Appli
36	40	49.4	433	4	US-08-979-847B-81	Sequence 81, Appli
37	39.5	48.8	89	4	US-09-252-991A-23279	Sequence 23279, A
38	39	48.1	86	4	US-08-311-731A-242	Sequence 242, App
39	39	48.1	114	4	US-09-252-991A-25913	Sequence 25913, A
40	39	48.1	145	4	US-09-482-273-240	Sequence 240, App
41	39	48.1	561	4	US-09-252-991A-24244	Sequence 24244, A
42	39	48.1	681	4	US-09-252-991A-24567	Sequence 24567, A
43	38.5	47.5	127	4	US-09-252-991A-20068	Sequence 20068, A
44	38.5	47.5	396	4	US-09-252-991A-28645	Sequence 28645, A
45	38.5	47.5	773	4	US-09-252-991A-31403	Sequence 31403, A

ALIGNMENTS

RESULT :
US-09-252-991A-32674
; Sequence 32674, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 10796.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32674
; LENGTH: 603
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32674

Query Match 59.9%; Score 48.5; DB 4; Length 603;
Best Local Similarity 42.9%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 3; Indels 9; Gaps 1;

QY 1 THRPPM-----WSPW 12
Db 576 TRPPRTAFAGSGMSPAW 596

RESULT 2
US-09-252-991A-32732
; Sequence 32732, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32732
; LENGTH: 521
; TYPE: PRT

! ORGENISM: Pseudomonas aeruginosa
US-09-252-991A-32732

Query Match 58.0%; Score 47; DB 4; Length 521;
Best Local Similarity 50.0%; Pred. No. 29;
Matches 7; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 3 RPPMSP-----VKF 12
|||:|:|:|:
DB 321 RPPLHPAVALIWP 334

RESULT 3
US-09-252-991A-31281
; Sequence 31281, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31281
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31281

Query Match 55.6%; Score 45; DB 4; Length 145;
Best Local Similarity 63.6%; Pred. No. 16;
Matches 7; Conservative 0; Mismatches 4; Indels 4; Gaps 0;

QY 1 THRPMPMSPVW 11
|||:|:|:|:
DB 128 TRPPPAWSANW 138

RESULT 4
US-09-618-425-2
; Sequence 2, Application: US/09618425
; Patent No. 6475744
; GENERAL INFORMATION:
; APPLICANT: Reppert, Steven M.
; APPLICANT: Weaver, David R.
; APPLICANT: Zylka, Mark
; APPLICANT: Jin, Xiaowei
; APPLICANT: Kume, Kazuhiko
; APPLICANT: Sirram, Sathyanarayanan
; APPLICANT: Shearman, Lauren
; TITLE OF INVENTION: METHODS FOR IDENTIFYING COMPOUNDS WHICH
; FILE REFERENCE: 00786-428001
; CURRENT APPLICATION NUMBER: US/09/618,425
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/203,005
; PRIOR FILING DATE: 2000-05-10
; PRIOR APPLICATION NUMBER: 60/145,363
; PRIOR FILING DATE: 1999-07-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 1197
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-618-425-2

Query Match 54.3%; Score 44; DB 4; Length 1197;

Best Local Similarity 66.7%; Pred. No. 1.7e-02;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 THRPMPMSP 9
|||:|:|:|:
DB 810 SHRAPLWSP 818

RESULT 5
US-08-660-963-13
; Sequence 13, Application US/08660963
; Patent No. 5852187
; GENERAL INFORMATION:
; APPLICANT: Thorne, Michael O.
; APPLICANT: Gaylinn, Bruce D.
; APPLICANT: Horikawa, Reiko
; APPLICANT: Lyons Jr., Charles E.
; TITLE OF INVENTION: MOLECULAR CLONING OF THE OVINE PITUITARY
; TITLE OF INVENTION: GROWTH HORMONE RELEASING HORMONE RECEPTOR
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: POPHAM, HAIR, SCHNOBRICH & KAUFMAN, LTD.
; STREET: Metropolitan Square Building, Suite 800, 1450
; STREET: G. Street
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/660,963
; FILING DATE: 12-JUN-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: O'Shaughnessy, Brian P.
; REGISTRATION NUMBER: 32,747
; REFERENCE/DOCKET NUMBER: 18046.036
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-824-8000
; TELEFAX: 202-824-8199
; TELEX: 248516
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 501 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-660-963-13

Query Match 53.1%; Score 43; DB 2; Length 501;
Best Local Similarity 54.5%; Pred. No. 1e+02;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 THRPMPMSPVW 11
|||:|:|:|:
DB 331 THRPPLCWGELW 341

RESULT 6
US-09-252-991A-20271
; Sequence 20271, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A

```
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20271
; LENGTH: 110
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20271

Query Match          52.5%; Score 42.5; DB 4; Length 110;
Best Local Similarity 77.8%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 4 PPMSPVWP 12
   ||| |||
Db 57 PPRW-PVWP 64

RESULT 7
US-09-252-991A-25448
; Sequence 25448, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25448
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25448

Query Match          51.9%; Score 42; DB 4; Length 345;
Best Local Similarity 58.3%; Pred. No. 95;
Matches 7; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 3 RPPXMS--PVWP 12
   ||| |||
Db 87 RPPSWTRWPPWP 98

RESULT 8
US-09-252-991A-28965
; Sequence 28965, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28965
; LENGTH: 441
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
```

```
; NAME/KEY: UNSURE
; LOCATION: (338), (413)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-28965

Query Match          51.9%; Score 42; DB 4; Length 441;
Best Local Similarity 58.3%; Pred. No. 1.2e+02;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 THRPWMSPVWP 12
   ||| |||
Db 414 TSRPPGWSAITP 425

RESULT 9
US-07-731-157A-2
; Sequence 2, Application US/07731157A
; Patent No. 5457032
; GENERAL INFORMATION:
; APPLICANT: Quax, Wilhelmus J.
; APPLICANT: Misset, Onno
; APPLICANT: Van der Laan, Jan M.
; APPLICANT: Lenting, Herman B.M.
; TITLE OF INVENTION: Mutated beta-lactam acylase genes
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: COOLEY GODWARD CASTRO HUDDLESON & TATUM
; STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (EPC)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/731,157A
; FILING DATE: 19910509
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 90200962
; FILING DATE: 18-APR-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: RAE-VENTER P.R.D., BARBARA
; REGISTRATION NUMBER: 32,750
; REFERENCE/DOCKET NUMBER: GBRO-027/00US
; TELEPHONE: 415-494-7622
; TELEFAX: 415-857-0663
; TELEX: 380816 COOLEY PA
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 720 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-731-157A-2

Query Match          51.9%; Score 42; DB 1; Length 720;
Best Local Similarity 62.5%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 PPMSPVWP 12
   ||| |||
Db 444 PPMTPWP 451

RESULT 10
US-08-541-780-2
; Sequence 2, Application US/08541780
; Patent No. 5935831
```

; GENERAL INFORMATION:
 ; APPLICANT: Quax, Wilhelmus J.
 ; APPLICANT: Misset, Onno
 ; APPLICANT: Van der Laan, Jan M.
 ; APPLICANT: Lenting, Herman B.M.
 ; TITLE OF INVENTION: Mutated beta-lactam acylase genes
 ; NUMBER OF SEQUENCES: 50
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: COOLEY GOWARD CASTRO HUDDLESON & TATUM
 ; STREET: FIVE PALO ALTO SQUARE, 4TH FLOOR
 ; CITY: PALO ALTO
 ; STATE: CALIFORNIA
 ; COUNTRY: USA
 ; ZIP: 94306
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25 (EPO)
 ; CURRENT APPLICATION DATA: US/08/541,780
 ; APPLICATION NUMBER: US/08/541,780
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/731,157
 ; FILING DATE:
 ; APPLICATION NUMBER: EP 90200962
 ; FILING DATE: 18-APR-1990
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: RAE-VENTER PH.D., BARBARA
 ; REGISTRATION NUMBER: 32,750
 ; REFERENCE/DOCKET NUMBER: G8RO-027/000US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 415-494-7622
 ; TELEFAX: 415-857-0663
 ; TELEX: 380816 COOLEY PA
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 720 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-541-780-2

Query Match 51.9%; Score 42; DB 2; Length 720;
 Best Local Similarity 62.5%; Pred. No. 2e+02;
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 5 PMWSPWVP 12
 Db 444 PWTFTWP 451

RESULT 11
 ; Sequence 30596, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 30596
 ; LENGTH: 469
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-30596

Query Match 51.2%; Score 41.5; DB 4; Length 469;
 Best Local Similarity 43.8%; Pred. No. 1.5e+02;
 Matches 7; Conservative 1; Mismatches 1; Indels 7; Gaps 1;

Qy 4 PMWSPV-----WP 12
 Db 105 PVMSPACATTARCWP 120

RESULT 12
 ; Sequence 24619, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 24619
 ; LENGTH: 882
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-24619

Query Match 51.2%; Score 41.5; DB 4; Length 882;
 Best Local Similarity 60.0%; Pred. No. 2.8e+02;
 Matches 6; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

Qy 4 PPM-WSPWVP 12
 Db 870 PPLAWPLWVP 879

RESULT 13
 ; Sequence 22258, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 22258
 ; LENGTH: 160
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-252-991A-22258

Query Match 50.6%; Score 41; DB 4; Length 160;
 Best Local Similarity 60.0%; Pred. No. 62;
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 THRPWMSPV 10
 Db 97 TFRPPSWKPL 106

RESULT 14
 US-09-252-991A-18776
 ; Sequence 18776, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252.991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/374,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 18776
 ; LENGTH: 266
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-18776

Query Match 50.6%; Score 41; DB 4; Length 266;
 Best Local Similarity 77.8%; Pred. No. 1e-02;
 Matches 7; Conservative 0; Mismatches 2; Indels 3; Gaps 0;

QY 1 THREPMKSP 9
 Db 121 TPRPGWSP 129

RESULT 15
 US-09-786-240-31
 ; Sequence 31, Application US/09786240
 ; Patent No. 6558935
 ; GENERAL INFORMATION:
 ; APPLICANT: INCYTE PHARMACEUTICALS, INC.
 ; APPLICANT: TANG, Y. Tom
 ; APPLICANT: CORLEY, Neil C.
 ; APPLICANT: GUEGLER, Karl J.
 ; APPLICANT: BAUGHN, Mariah R.
 ; APPLICANT: LAL, Preeti
 ; APPLICANT: YUE, Henry
 ; APPLICANT: HILLMAN, Jennifer L.
 ; APPLICANT: AZIMZAI, Yaida
 ; TITLE OF INVENTION: HUMAN TRANSFERASE PROTEINS
 ; FILE REFERENCE: PF-0392.PCT
 ; CURRENT APPLICATION NUMBER: US/09/786.240
 ; CURRENT FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 09/150,657; unassigned: 09/186,779; unassigned: 60/133,642
 ; PRIOR FILING DATE: 1998-09-10; 1998-09-10; 1998-11-04; 1998-11-04; 1999-05-11
 ; NUMBER OF SEQ ID NOS: 33
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 31
 ; LENGTH: 478
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; PUBLICATION INFORMATION:
 ; DATABASE ACCESSION NUMBER: g24438.4
 US-09-786-240-31

Query Match 50.6%; Score 41; DB 4; Length 478;
 Best Local Similarity 66.7%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 3 RPPVMSFVW 11
 Db 184 RPPGWLPCW 192

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 20:50:36 ; Search time 39 seconds
(without alignments)
29.590 Million cell updates/sec

Title: US-09-995-804A-2

Perfect score: 81

Sequence: 1 THREPMWSPVWP 12

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 1580

Minimum DB seq length: 0

Maximum DB seq length: 12

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 76:*

1: PIR1:*

2: PIR2:*

3: PIR3:*

4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	48.1	10	2 S39030	lysyl-bradykinin -
2	31	38.3	9	2 A61081	tryptophyllin, bas
3	28	34.6	9	2 A61363	bradykinin - commo
4	28	34.6	9	2 S65433	bradykinin - horn
5	28	34.6	9	2 A43065	hydroxyproline-3-b
6	28	34.6	11	2 A61365	phyllokinin - Rohd
7	28	34.6	11	2 S13279	lie-ser-bradykinin
8	28	34.6	12	2 A61360	vespakinin M - hor
9	28	34.6	12	2 A61359	vespakinin X - hor
10	26	32.1	9	2 S60246	critho-kinin - ch
11	25	30.9	9	2 A61057	Thr-6 bradykinin -
12	25	30.9	9	2 A26744	bradykinin-like pe
13	25	30.9	9	2 A60579	bradykinin-like pe
14	25	30.9	11	2 B26744	megascollakinin -
15	24	29.6	8	2 A41117	acetylcholinestera
16	24	29.6	12	2 P00046	ATP synthase D cha
17	23	28.4	11	2 P00273	IG heavy chain CRD
18	23	28.4	12	2 B39600	neural cell adhesi
19	22	27.2	7	2 P00098	glucuronosyltransf
20	22	27.2	7	4 A58725	virucocin - destro
21	22	27.2	11	2 P00731	unidentified 5.7/3
22	21	25.9	8	2 S15422	adipokinetic hormo
23	21	25.9	8	2 A58641	adipokinetic hormo
24	21	25.9	8	2 A39338	glycine reductase
25	21	25.9	8	2 S21663	neuropeptide - fio
26	21	25.9	10	1 RHAQ2	gonadoliberin II -
27	21	25.9	10	1 A61126	gonadoliberin - sp
28	21	25.9	10	2 B46C30	gonadoliberin II -
29	21	25.9	10	2 S18396	probable glucose-6

ALIGNMENTS

RESULT 1

S39030

lysyl-bradykinin - rainbow trout

C:Species: Oncorhynchus mykiss (rainbow trout)

C:Date: 19-May-1994 #sequence_revision 19-Apr-1996 #text_change 18-Aug-2000

C:Accession: S39030

R:Conlon, J.M.; Olson, K.R.

FEBS Lett. 334, 75-78, 1993

A:Title: Purification of a vasoactive peptide related to lysyl-bradykinin from trout

A:Reference number: S39030; MUID:94039817; PMID:8224232

A:Accession: S39030

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-10 <CON>

C:Superfamily: unassigned animal peptides

Query Match 48.1%; Score 39; DB 2; Length 10;

Best Local Similarity 75.0%; Pred. No. 3.9;

Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWSVP 10

DB 2 RPPGWSPL 9

RESULT 2

A61081

tryptophyllin, basic - Rohde's leaf frog

C:Species: Phyllomedusa rohdei (Rohde's leaf frog)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 18-Aug-2000

C:Accession: A61081

R:Montecucchi, P.C.; Vincenti, M.; Lazzarini, A.M.; Rusconi, L.; Erspamer, V.

Int. J. Pept. Protein Res. 33, 391-395, 1989

A:Title: Isolation, structure determination and synthesis of a novel tryptophan-conta

A:Reference number: A61081

A:Accession: A61081

A:Molecule type: protein

A:Residues: 1-7 <MON>

C:Comment: The biological activity of this peptide was not determined.

C:Superfamily: unassigned animal peptides

C:Keywords: amidated carboxyl end; hydroxyproline; skin

F:3/Modified site: 4-hydroxyproline (Pro) #status experimental

F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 38.3%; Score 31; DB 2; Length 7;

Best Local Similarity 57.1%; Pred. No. 2.8e+05;

Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 RPPMWSVP 9

DB 1 KPPSWIP 7

```

RESULT 3
A61363
bradykinin - common frog
C:Species: Rana temporaria (common frog)
C>Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Aug-2000
C:Accession: A61363
R:Anastasi, A.; Erspamer, V.; Bertaccini, G.
Comp. Biochem. Physiol. A 14, 43-52, 1965
A:Title: Occurrence of bradykinin in the skin of Rana temporaria.
A:Reference number: A61363
A:Accession: A61363
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <ANA>
C:Superfamily: unassigned animal peptides
C:Keywords: skin

Query Match 34.6%; Score 28; DB 2; Length 9;
Best Local Similarity 71.4%; Pred. No. 2.8e+05;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWSP 9
DB 1 RPPGFSP 7

RESULT 4
S65433
bradykinin - horn fly (fragment)
C:Species: Haematobia irritans (horn fly)
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 13-Mar-1997
C:Accession: S65433
R:Wijffels, G.; Fitzgerald, C.; Gough, J.; Riding, G.; Elvin, C.; Kemp, D.; Willadsen, E.
Eur. J. Biochem. 237, 414-423, 1996
A:Title: Cloning and characterization of angiotensin-converting enzyme from the dipteran
A:Reference number: S65431; MUID:96215437; PMID:8647080
A:Accession: S65433
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <WIF>
A>Note: the source is designated as Haematobia irritans exigua

Query Match 34.6%; Score 28; DB 2; Length 9;
Best Local Similarity 71.4%; Pred. No. 2.8e+05;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWSP 9
DB 1 RPPGFSP 7

RESULT 5
A43065
hydroxyproline-3-bradykinin - frog (Helleophryne purcelli)
C:Species: Helleophryne purcelli
C>Date: 07-Oct-1994 #sequence_revision 07-Oct-1994 #text_change 07-May-1999
C:Accession: A43065
R:Nakajima, T.; Yasuhara, T.; Erspamer, G.F.; Visser, J.
Experientia 35, 1133, 1979
A:Title: Occurrence of Hyp(3)-bradykinin in methanol extracts of the skin of the South A
A:Reference number: A43065; MUID:80024576; PMID:488255
A:Accession: A43065
A:Molecule type: protein
A:Residues: 1-9 <NAK>
C:Keywords: bradykinin; hydroxyproline; skin
F:3/Modified site: hydroxyproline (Pro) #status experimental

Query Match 34.6%; Score 28; DB 2; Length 9;
Best Local Similarity 71.4%; Pred. No. 2.8e+05;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWSP 9

RESULT 6
A61365
Phyllokinin - Rohde's leaf frog
N:Alternate names: bradykinyl-isoleucyl-tyrosine O-sulfate
C:Species: Phyllomedusa rohdei (Rohde's leaf frog)
C>Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Aug-2000
C:Accession: A61365
R:Anastasi, A.; Bertaccini, G.; Erspamer, V.
Br. J. Pharmacol. 27, 479-485, 1966
A:Title: Pharmacological data on phyllokinin (bradykinyl-isoleucyl-tyrosine O-sulfate)
A:Reference number: A61365; MUID:67179312; PMID:5970899
A:Accession: A61365
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-11 <ANA>
C:Superfamily: unassigned animal peptides
C:Keywords: sulfoprotein
F:11/Binding site: sulfate (Tyr) (covalent) #status experimental

Query Match 34.6%; Score 28; DB 2; Length 11;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWSP 9
DB 1 RPPGFSP 7

RESULT 7
S13279
Ile-Ser-bradykinin - human (fragment)
N:Alternate names: T-kinin
C:Species: Homo sapiens (man)
C>Date: 02-Dec-1993 #sequence_revision 13-Mar-1997 #text_change 24-Jul-1998
C:Accession: S13279
R:Wunderer, G.; Walter, I.; Eschenbacher, B.; Lang, M.; Kellermann, J.; Kindermann, G.
Biol. Chem. Hoppe-Seyler 371, 977-981, 1990
A:Title: Ile-Ser-bradykinin is an aberrant permeability factor in various human malign
A:Reference number: S13279; MUID:91166748; PMID:2076202
A:Accession: S13279
A:Molecule type: protein
A:Residues: 1-11 <WUN>
C:Keywords: bradykinin

Query Match 34.6%; Score 28; DB 2; Length 11;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWSP 9
DB 3 RPPGFSP 9

RESULT 8
A61360
vespakinin M - hornet (Vespa mandarinia)
C:Species: Vespa mandarinia
C>Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Aug-2000
C:Accession: A61360
R:Kishimura, H.; Yasuhara, T.; Yoshida, H.; Nakajima, T.
Chem. Pharm. Bull. 24, 2896-2897, 1976
A:Title: Vespakinin-M, a novel bradykinin analogue containing hydroxyproline, in the
A:Reference number: A61360; MUID:77114342; PMID:1017116
A:Accession: A61360
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-12 <KIS>
C:Superfamily: unassigned animal peptides
C:Keywords: hydroxyproline; venom
```


F;4/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match 34.6%; Score 28; DB 2; Length 12;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWSP 9
||| :||
Db 2 RPPGFSP 8

RESULT 9

A61359
vespakinin X - hornet (Vespa xanthoptera)
C:Species: Vespa xanthoptera
C>Date: 03-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 18-Aug-2000
C:Accession: A61359
R.Yasuhara, T.; Yoshida, H.; Nakajima, T.
Chem. Pharm. Bull. 25, 936-941, 1977
A:Title: Chemical investigation of the hornet (Vespa xanthoptera Cameron) venom. The structure of vespakinin X.
A:Reference number: A61359; MUID:97137853; PMID:284186
A:Accession: A61359
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-12 <VAS>
C:Superfamily: unassigned animal peptides
C:Keywords: venom

Query Match 34.6%; Score 28; DB 2; Length 12;
Best Local Similarity 71.4%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWSP 9
||| :||
Db 2 RPPGFSP 8

RESULT 10

B60246
ornitho-kinin - chicken
C:Species: Gallus gallus (chicken)
C>Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 18-Aug-2000
C:Accession: B60246
R.Kimura, M.; Sueyoshi, T.; Morita, T.; Tanaka, K.; Iwanaga, S.
Adv. Exp. Med. Biol. 247A, 359-367, 1989
A:Title: Ornitho-kininogen and ornitho-kinin: isolation, characterization and chemical structure.
A:Reference number: A60246; MUID:90102072; PMID:2603803
A:Accession: B60246
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-9 <KIMS>
C:Superfamily: unassigned animal peptides

Query Match 32.1%; Score 26; DB 2; Length 9;
Best Local Similarity 50.6%; Pred. No. 2.8e+05;
Matches 4; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWSPV 10
||| :||
Db 1 RPPGFSP 8

RESULT 11

A61057
Thr-6 bradykinin - scollid wasp (Colpa interrupta)
C:Species: Colpa interrupta
C>Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 18-Aug-2000
C:Accession: A61057
R.Piek, T.; Hue, B.; Mantel, P.; Nakajima, T.; Pelhate, M.; Yasuhara, T.
Comp. Biochem. Physiol. C 96, 157-162, 1990
A:Title: Threonine(6)-bradykinin in the venom of the wasp Colpa interrupta (F.) presynaptic release of bradykinin.
A:Reference number: A61057; MUID:91130217; PMID:1980872
A:Accession: A61057

A:Molecule type: protein

A:Residues: 1-9 <PIE>

C:Superfamily: unassigned animal peptides

C:Keywords: bradykinin; presynaptic neurotoxin; venom

Query Match 30.9%; Score 25; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 2.8e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWSP 9
||| :||
Db 1 RPPGFSP 7

RESULT 12

A26744
bradykinin-like peptide - garden dagger wasp
N:Alternate names: Thr-6-bradykinin
C:Species: Megascolia flavifrons (garden dagger wasp)
C>Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 18-Aug-2000
C:Accession: A26744
R.Yasuhara, T.; Mantel, P.; Nakajima, T.; Piek, T.
Toxicol. 25, 527-535, 1987
A:Title: Two kinins isolated from an extract of the venom reservoirs of the solitary wasp Megascolia flavifrons (Gardner).
A:Reference number: A94322; MUID:87293024; PMID:3617088
A:Accession: A26744
A:Molecule type: protein
A:Residues: 1-9 <VAS>
C:Superfamily: unassigned animal peptides

Query Match 30.9%; Score 25; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 2.8e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWSP 9
||| :||
Db 1 RPPGFSP 7

RESULT 13

A60579
bradykinin-like peptide - slider turtle
C:Species: pseudemys scripta (slider)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 18-Aug-2000
C:Accession: A60579
R.Conlon, J.M.; Hicks, J.W.; Smith, D.D.
Endocrinology 126, 985-991, 1990
A:Title: Isolation and biological activity of a novel kinin ((Thr(6))bradykinin) from the slider turtle.
A:Reference number: A60579; MUID:90126625; PMID:2298179
A:Accession: A60579
A:Molecule type: protein
A:Residues: 1-9 <CON>
C:Comment: This peptide increases aortic blood flow but, unlike bradykinin in mammalia
C:Superfamily: unassigned animal peptides
C:Keywords: plasma

Query Match 30.9%; Score 25; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 2.8e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWSP 9
||| :||
Db 1 RPPGFSP 7

RESULT 14

B26744
megascoliakinin - garden dagger wasp
N:Alternate names: 6-Thr-bradykinin-Lys-Ala
C:Species: Megascolia flavifrons (garden dagger wasp)
C>Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 18-Aug-2000
C:Accession: B26744; A28609
R.Yasuhara, T.; Mantel, P.; Nakajima, T.; Piek, T.

TOXICON 25, 527-535, 1987
 A:Title: Two kinins isolated from an extract of the venom reservoirs of the solitary wasp
 A:Reference number: A94322; MJID:87293024; PMID:3617089
 A:Accession: B26744
 A:Molecule type: protein
 A:Residues: 1-11 <YAS>
 R:Nakajima, T.; Diek, T.; Yashuara, T.; Mantel, P.
 Toxicon 26, 34, 1988
 A:Title: Two kinins isolated from the venom of Megascolia flavifrons.
 A:Reference number: A28609
 A:Accession: A28609
 A:Molecule type: protein
 A:Residues: 1-11 <NAK>
 C:Superfamily: unassigned animal peptides
 C:Keywords: bradykinin; presynaptic neurotoxin; venom

Query Match 30.9%; Score 25; DS 2; Length 11;
 Best Local Similarity 57.1%; Pred.No. 5.1e+02;
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Cy 3 RPPMWSP 9
 ||| :|
 Db 1 RPPGFP 7

RESULT 15
 A4117
 acetylcholinesterase (EC 3.1.1.7), venom - Asian cobra (fragment)
 C:Species: Naja naja oxiana (Asian cobra, Oxus cobra)
 C:Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 23-Jun-1993
 C:Accession: A4117
 R:Kreienkamp, H.J.; Weise, C.; Raba, R.; Aaviksaar, A.; Hucho, F.
 Proc. Natl. Acad. Sci. U.S.A. 88, 6117-6121, 1991
 A:Title: Anionic subsites of the catalytic center of acetylcholinesterase from Torpedo a
 A:Reference number: A4117; MCIID:91296772; PMID:2068091
 A:Accession: A4117
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-8 <KRE>
 C:Keywords: carboxylic ester hydrolase

Query Match 29.6%; Score 24; DS 2; Length 8;
 Best Local Similarity 75.0%; Pred.No. 2.8e+05;
 Matches 3; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 6 WSP 9
 ||| :|
 Db 4 MWNP 7

Search completed: October 22, 2003, 20:54:55
 Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 20:48:56 ; Search time 22 Seconds
(without alignments)
25.651 Million cell updates/sec

Title: US-09-995-804A-2
Perfect score: 81
Sequence: 1 THRPWMSPVMP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 501

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	39	48.1	10	1 BRK ONCMY	Q9PRZL: oncorhynch
2	31	38.3	7	1 TPFY PACDA	P83455 pachymedusa
3	25	30.9	9	1 KNL3 BOMVA	P83558 bombina var
4	25	30.9	11	1 BRK MEGFL	P12797 megascolia
5	23	28.4	12	1 RFL CONSP	P58805 conus spuri
6	22	27.2	10	1 AKH2 LOEMI	P81626 locusta mig
7	21	25.9	6	1 EIG1 LITRU	P82096 litorea rub
8	21	25.9	8	1 AKH MELM	P25423 melolontha
9	21	25.9	10	1 GON2 CHICK	P37043 gallus gall
10	21	25.9	10	1 GON3 PETMA	P10948 petromyzon
11	21	25.9	10	1 GON1 SQUAC	P27429 squallus aca
12	20	24.7	8	1 HTF1 PERAM	P04548 periplaneta
13	20	24.7	8	1 HTF2 TENMO	P25419 tenebrio mo
14	20	24.7	9	1 RPCH PANBO	P08939 pardalio bo
15	20	24.7	10	1 GON3 ONCKE	P20367 oncorhynch
16	20	24.7	10	1 HTF NAUCI	P10939 nauphoeta c
17	19.5	24.1	11	1 LPW TETH	P05624 thermus the
18	19	23.5	5	1 BPP7 BOTIN	P30425 bothriops ir
19	19	23.5	11	1 BPP AGKRP	P04562 agkistrodon
20	19	23.5	11	1 RANC RANPI	P08951 rana pipien
21	18.5	22.8	9	1 COW CONVE	P83047 conus ventr
22	18	22.2	8	1 AKH LIBAU	P25418 libellula a
23	18	22.2	10	1 BPP2 BOTIN	P10422 bothriops ja
24	18	22.2	10	1 BPP7 BOTJA	P01022 bothriops ja
25	18	22.2	10	1 BPP VEPAS	P31351 vipera aspi
26	18	22.2	10	1 GON1 PETMA	P04378 petromyzon
27	18	22.2	12	1 UP01 CAEEL	P55954 caenorhabdi
28	17.5	21.6	10	1 LABA JATMU	P13270 jatropa mu
29	17	21.0	8	1 AKH TABAT	P14595 tabanus atr
30	17	21.0	8	1 HTF2 PERAM	P04549 periplaneta
31	17	21.0	10	1 HTF1 ROMMI	P18110 romalea mic
32	17	21.0	10	1 HTF2 CARMO	P11385 carausus m
33	17	21.0	10	1 HTF TABAT	P14596 tabanus atr

```

34 17 21.0 11 1 CEPI ACHFU      P22790 achatina fu
35 17 21.0 11 1 MORN_HUMAN    P01163 homo sapien
36 16 19.8 7 1 UF04_MOUSE     P38642 mus musculu
37 16 19.8 9 1 UPA3_HUMAN     P30089 homo sapien
38 16 19.8 12 1 FAR7_PENMO    P83322 penaeus mor
39 16 19.8 12 1 TM2A_METMA    P80652 methanosarc
40 15 19.1 12 1 PK4_PERAM      P82619 periplaneta
41 15 18.5 5 1 UF01_MOUSE     P38639 mus musculu
42 15 18.5 7 1 WWA1_ACHFU      P35919 achatina fu
43 15 18.5 7 1 WWA2_ACHFU      P35920 achatina fu
44 15 18.5 7 1 WWA3_ACHFU      P35921 achatina fu
45 15 18.5 9 1 NEF_HV128      P12481 human immun

```

ALIGNMENTS

```

RESULT 1
BRK ONCMY
AC Q9PRZL STANDARD; PRT; 10 AA.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lysyl-bradykinin-like.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE.
RX MEDLINE=94339817; PubMed=8224232;
RA Conlon J.W., Olson K.R.;
RT "Purification of a vasoactive peptide related to lysyl-bradykinin from
RT trout plasma.";
RL FEBS Lett. 334:75-78(1993).
CC -!- FUNCTION: SMOOTH MUSCLE CONTRACTION. PROBABLY PLAYS A ROLE FOR
CC -!- THIS SYSTEM IN CARIOVASCULAR REGULATION IN FISH.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Plasma.
CC -!- SIMILARITY: BELONGS TO THE BRADYKININ FAMILY.
DR PIR: S39030; S39030.
KW Bradykinin; Vasodilator.
SQ SEQUENCE 10 AA; 1193 MW; 33C59075A3786777 CRC64;

```

Query Match 48.1%; Score 39; DB 1; Length 10;
Best Local Similarity 75.0%; Pred. No. 2.3;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 3 RPPWMSPV 10
DB |||||
  2 RPPGWSP 9

```

```

RESULT 2
TPFY PACDA
ID TPFY_PACDA STANDARD; PRT; 7 AA.
AC P83455;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Trypophyllin-1 (pdt-1).
OS Pachymedusa dactinolor (Giant mexican leaf frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonidae; Hylidae;
OC Phyllomedusinae; Pachymedusa.
OX NCBI_TaxID=75988;
RN [1]
RP SEQUENCE, MASS SPECTROMETRY, HYDROXYLATION OF PRO-3, AND AMIDATION OF
RP PRO-7.
RC TISSUE=Skin secretion;
RA Chen T.B., Orr D.F., Shaw C.;

```

```

RT "Pachymedusa dactylophylloides" (PDR-1): structural
RT characterization, pharmacological activity and cloning of precursor
RT cDNA.
RL Submitted (SEP-2002): to the SWISS-PROT data bank.
CC -!- FUNCTION: Viscerotropic. Has selective relaxing activity on vascular
CC smooth muscle.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- MASS SPECTROMETRY: MW=809.2; METHOD=MALDI.
DR GO: GO:0005576; C:extracellular; NAS.
DR GO: GO:0045986; P:negative regulation of smooth muscle contractility; NAS.
KW Amphibian defense peptide; Amidation; Hydroxylation.
FT MOD_RES 3 3 HYDROXYLATION.
FT MOD_RES 7 7 AMIDATION.
SQ SEQUENCE 7 AA; 794 MW; 7772D37DC7776350 CRC64;

Query Match 38.3%; Score 31; DB 1; Length 7;
Best Local Similarity 57.1%; Pred. No. 1.3e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 RPPMWP 9
DB 1 KPPAWP 7

RESULT 3
ID KNL3_BOMVA STANDARD; PRT; 9 AA.
AC P81059;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE [Thr6]bradykinin.
OS Bombina variegata (Yellow-bellied toad).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Archeobatrachia; Bombinatoridae; Bombina.
OX NCBI_TaxID=8348;
RN [1]
RP SEQUENCE, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC TISSUE=Skin secretion;
RA Chen T.B., Orr D.F., Boursion A.J., McKean S., Rao P.F., Shaw C.;
RT "Cloning and post-translational processing of frog skin kininogens.";
RL Submitted (JUL-2001): to the SWISS-PROT data bank.
CC -!- FUNCTION: [Thr6]bradykinin produces in vitro relaxation of rat
CC arterial smooth muscle and constriction of intestinal smooth
CC muscle.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
CC -!- SIMILARITY: BELONGS TO THE BRADYKININ FAMILY.
KW Amphibian defense peptide; Vasodilator; Bradykinin.
SQ SEQUENCE 9 AA; 1074 MW; 3393D771A9C86777 CRC64;

Query Match 30.9%; Score 25; DB 1; Length 9;
Best Local Similarity 57.1%; Pred. No. 1.3e+05;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWP 9
DB 1 RPPGFP 7

RESULT 4
ID BRK_MEGFL STANDARD; PRT; 11 AA.
AC P12797;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Megascolinin ([Thr6]bradykinin-Lys-Ala) [Contains: Bradykinin-like
DE peptide ([Thr6]bradykinin)].
OS Megascolia flavifrons (Garden dagger wasp) (Solitary wasp).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Scolioidea;

```

```

OC Scolioidea; Megascolia.
OX NCBI_TaxID=7437;
RN [1]
RP SEQUENCE.
RC TISSUE=Venom;
RX MEDLINE=87293024; PubMed=3617088;
RA Yashuara T., Mantel P., Nakajima T., Piek T.;
RT "Two kinins isolated from an extract of the venom reservoirs of the
RT solitary wasp Megascolia flavifrons.";
RL Toxicon 25:527-535(1987).
RN [2]
RP SEQUENCE.
RC TISSUE=Venom;
RA Nakajima T., Piek T., Yashuara T., Mantel P.;
RT "Two kinins isolated from the venom of Megascolia flavifrons.";
RL Toxicon 26:34-34(1988).
CC -!- FUNCTION: Both proteins have bradykinin-like, although lower
CC activities (e.g. smooth muscle contraction).
CC -!- SUBCELLULAR LOCATION: Secreted; wasp venom reservoirs.
CC -!- SIMILARITY: BELONGS TO THE BRADYKININ FAMILY.
DR PIR; B26744; B26744.
KW Bradykinin; Vasodilator.
FT PEPTIDE 1 11 MEGASCOLIAKININ.
FT PEPTIDE 1 9 BRADYKININ-LIKE PEPTIDE.
SQ SEQUENCE 11 AA; 1273 MW; 33867393D771A9C8 CRC64;

Query Match 30.9%; Score 25; DB 1; Length 11;
Best Local Similarity 57.1%; Pred. No. 2.8e+02;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMWP 9
DB 1 RPPGFP 7

RESULT 5
ID RFL_CONSP STANDARD; PRT; 12 AA.
AC P58305;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Conorfamide-Srl.
OS Conus spurius (Alphabet cone).
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC Neogastropoda; Conoidea; Conidae; Conus.
OX NCBI_TaxID=192919;
RN [1]
RP SEQUENCE, SYNTHESIS, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=21605839; PubMed=11738233;
RA Maillo M., Aguilar M.B., Lopez-Vera E., Craig A.G., Bula G.;
RA Olivera B.M., Heimer de la Cotera E.P.;
RT "Conorfamide, a Conus venom peptide belonging to the Rfamamide family of
RT neuropeptides.";
RL Toxicon 40:401-407(2002).
CC -!- FUNCTION: Causes hyperactivity in mice.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the venom duct.
CC -!- MASS SPECTROMETRY: MW=1454.8; METHOD=Electrospray.
CC -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neurotoxin; Toxin; Amidation.
FT MOD_RES 12 12 AMIDATION.
SQ SEQUENCE 12 AA; 1456 MW; 2510671E49D772D3 CRC64;

Query Match 28.4%; Score 23; DB 1; Length 12;
Best Local Similarity 50.0%; Pred. No. 5.9e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 RPPMWPVW 11
DB 1 RPPGFP 7

```

```

Db          2 PMGWVPVF 9

RESULT 6
AKHX LOCMI
ID AKHX LOCMI STANDARD; PRT; 10 AA.
AC P81626;
DT 30-MAY-2003 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Peptide hormone.
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pserygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridocephala;
OC Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE.
RC TISSUE=Corpora cardiaca;
RA Siebert X.J.;
RL Submitted (DEC-1998) to the SWISS-PROT data bank.
CC -!- FUNCTION: PROBABLY INVOLVED IN THE REGULATION OF LOCUST
CC INTERMEDIARY METABOLISM, BEHAVIOR AND/OR DEVELOPMENT.
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR InterPro: IPR002647; AKH.
DR PROSITE: PS00256; AKH; 1.
KW Neuropeptide; Amidation; Pyroglutamate carboxylic acid.
FT MOD_RES 1 1 PYROGLUTAMATE CARBOXYLIC ACID.
FT MOD_RES 10 10 AMIDATION.
SQ SEQUENCE 10 AA; 1222 MW; 81BFF67AB415B9D1 CRC64;

Query Match 27.28; Score 22; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 6.9e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 WSP 9
|||
|||
Db 8 WSP 10

RESULT 7
EI01 LITRU
ID EI01 LITRU STANDARD; PRT; 6 AA.
AC P82036;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Electrin 1.
OS Litoria rubella (Desert tree frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC Pelodyadinae; Litoria.
OX NCBI_TaxID=104895;
RN [1]
RP SEQUENCE.
RC TISSUE=Skin secretion;
RA Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RL "Peptides from the skin glands of the Australian buzzing tree frog
RT Litoria electrica. Comparison with the skin peptides from Litoria
RT rubella."
RL Aust. J. Chem. 52:639-645(1999).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Skin.
KW Amphibian defense peptide; Amidation.
FT MOD_RES 6 6 AMIDATION.
SQ SEQUENCE 6 AA; 792 MW; 6683704772C9A00C CRC64;

Query Match 25.9%; Score 21; DB 1; Length 6;
Best Local Similarity 66.7%; Pred. No. 1.3e+05;
Matches 2; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 9 PVW 11
|||

RESULT 8
AKHX MELML
ID AKHX MELML STANDARD; PRT; 8 AA.
AC P25423;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Adipokinetic hormone (AKH).
OS Melolontha melolontha (Cockchafer).
OS Geotrupes stercorosus (Dor beetle), and
OS Pachnoda marginata (flower beetle).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pserygota;
OC Neoptera; Endopterygota; Coleoptera; Polyphaga; Scarabaeiformia;
OC Scarabaeidae; Melolonthinae; Melolontha.
OX NCBI_TaxID=7061, 7087, 7058;
RN [1]
RP SEQUENCE.
RC SPECIES=M.melolontha, and G.stercorosus; TISSUE=Corpora cardiaca;
RX MEDLINE=91248100; PubMed=2039445;
RA Gaede G.;
RL "A unique charged tyrosine-containing member of the adipokinetic
RT hormone/red-pigment-concentrating hormone peptide family isolated and
RT sequenced from two beetle species."
RL Biochem. J. 275:671-677(1991).
RN [2]
RP SEQUENCE.
RC SPECIES=P.marginata; TISSUE=Corpora cardiaca;
RX MEDLINE=92265187; PubMed=1586453;
RA Gaede G., Lopata A., Kellner R., Rinehart K.L. Jr.;
RL "Primary structures of neuropeptides isolated from the corpora
RT cardiaca of various ceterid beetle species determined by
RT pulsed-liquid phase sequencing and tandem fast atom bombardment mass
RT spectrometry."
RL Bio. Chem. Hoppe-Seyler 373:133-142(1992).
CC -!- FUNCTION: THIS HORMONE, RELEASED FROM CELLS IN THE CORPORA
CC CARDIACA AFTER THE BEGINNING OF FLIGHT, CAUSES RELEASE OF
CC DIGLYCERIDES FROM THE FAT BODY AND THEN STIMULATES THE FLIGHT
CC MUSCLES TO USE THESE DIGLYCERIDES AS AN ENERGY SOURCE.
CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.
DR PIR; A58641; A58641.
DR PIR; S15422; S15422.
DR InterPro: IPR002647; AKH.
DR PROSITE: PS00256; AKH; 1.
KW Neuropeptide; Amidation; Flight; Pyroglutamate carboxylic acid.
FT MOD_RES 1 1 PYROGLUTAMATE CARBOXYLIC ACID.
FT MOD_RES 8 8 AMIDATION.
SQ SEQUENCE 8 AA; 1022 MW; 867AB775AB544736 CRC64;

Query Match 25.9%; Score 21; DB 1; Length 8;
Best Local Similarity 60.0%; Pred. No. 1.3e+05;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 WSPVW 11
|||
|||
Db 4 WSPDW 8

RESULT 9
CON2 CHICK
ID CON2 CHICK STANDARD; PRT; 10 AA.
AC P37043; P20408; P81750;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Gonadoliberin II (Gonadotropin-releasing hormone II) (GrRH-II)
DE (LR-RH II) (Luliberin II).
OS Gallus gallus (Chicken).
OS Alligator mississippiensis (American alligator),
OS Squalus acanthias (Spiny dogfish),
OS Hydrolagus coliei (Spotted ratfish), and

```

OS *Clupea pallasii* (Pacific herring).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031, 8496, 7797, 7873, 30724;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=Chicken; TISSUE=Hypothalamus;
 RX MEDLINE=94222059; PubMed=6427779;
 RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
 RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
 RA Matsuo H.;
 RA "Identification of the second gonadotropin-releasing hormone in
 RT chicken hypothalamus: evidence that gonadotropin secretion is
 RT probably controlled by two distinct gonadotropin-releasing hormones
 RT in avian species.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:3874-3878 (1994).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=A. mississippiensis; TISSUE=Brain;
 RX MEDLINE=9335336; PubMed=862082;
 RA Lovejoy D.A., Fischer W.H., Parker D.B., McRory J.E., Park M.,
 RA Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
 RA "Primary structure of two forms of gonadotropin-releasing hormone
 RT from brains of the American alligator (*Alligator mississippiensis*).";
 RL Regul. Pept. 33:105-116 (1991).
 RN [3]
 RP SEQUENCE.
 RC SPECIES=S. acanthias; TISSUE=Brain;
 RX MEDLINE=9235300; PubMed=1631133;
 RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
 RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
 RA "Distinct sequence of gonadotropin-releasing hormone (GnRH) in
 RT dogfish brain provides insight into GnRH evolution.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377 (1992).
 RN [4]
 RP SEQUENCE.
 RC SPECIES=H. colliei; TISSUE=Brain;
 RX MEDLINE=91340067; PubMed=1678723;
 RA Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
 RA Lee T.;
 RA "Primary structure of gonadotropin-releasing hormone from the brain
 RT of a holoccephalan (ratfish: *Hydrolagus colliei*).";
 RL Gen. Comp. Endocrinol. 82:152-161 (1991).
 RN [5]
 RP SEQUENCE, AND FUNCTION.
 RC SPECIES=C. pallasii; TISSUE=Brain, and Pituitary;
 RX MEDLINE=20114351; PubMed=10650929;
 RA Carlsfeld J., Powell J.F.P., Park M., Fischer W.H., Craig A.G.,
 RA Chang J.P., Rivier J.E., Sherwood N.M.;
 RA "Primary structure and function of three gonadotropin-releasing
 RT hormones, including a novel form, from an ancient teleost, herring.";
 RL Endocrinology 141:505-512 (2000).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 DR PIR: A61126; A61126.
 DR PIR: B46030; B46030.
 DR PIR: B60066; B60066.
 DR InterPro: IPR002012; GnRH.
 DR Pfam: PF00446; GnRH; 1.
 DR PROSITE: PS00473; GnRH; 1.
 KW Hormone; Amidation; Hypothalamus; Pyrrolidone carboxylic acid.
 FT MOD_RES 10 10 AMIDATION.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1254 MW; 284B2E43787F5A3 CRC64;
 Query Match 25.9%; Score 21; DB 1; Length 10;
 Best Local Similarity 60.6%; Pred. No. 9.7e+02;
 Matches. 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 7 WSPVW 11
 DB 3 WSHGW 7

RESULT 10
 GON3_PETWA STANDARD; PRT; 10 AA.
 ID GON3_PETWA STANDARD; PRT; 10 AA.
 AC P30948;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GnRH-III);
 DE (Luliberin III).
 CS Petromyzon marinus (Sea lamprey).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
 CC Petromyzontiformes; Petromyzontidae; Petromyzon.
 CX NCBI_TaxID=7757;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=93178116; PubMed=8440174;
 RA Sower S.A., Chiang Y.-C., Lovas S., Conlon J.M.;
 RA "Primary structure and biological activity of a third gonadotropin-
 RT releasing hormone from lamprey brain.";
 RL Endocrinology 132:1125-113 (1993).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 DR InterPro: IPR002012; GnRH.
 DR Pfam: PF00446; GnRH; 1.
 DR PROSITE: PS00473; GnRH; 1.
 KW Hormone; Amidation; Hypothalamus; Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1277 MW; 284B36237AALF5A3 CRC64;
 Query Match 25.9%; Score 21; DB 1; Length 10;
 Best Local Similarity 60.6%; Pred. No. 9.7e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 7 WSPVW 11
 DB 3 WSHDW 7

RESULT 11
 GONL_SQUAC STANDARD; PRT; 10 AA.
 ID GONL_SQUAC STANDARD; PRT; 10 AA.
 AC P27429;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gonadoliberin (Gonadotropin-releasing hormone) (GnRH) (LH-RH)
 DE (Luliberin).
 CS Squalus acanthias (Spiny dogfish).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
 CC Elasmobranchii; Squalia; Squaloidei; Squalidae; Squalus.
 CX NCBI_TaxID=7797;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=92335300; PubMed=1631133;
 RA Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
 RA Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
 RA "Distinct sequence of gonadotropin-releasing hormone (GnRH) in
 RT dogfish brain provides insight into GnRH evolution.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377 (1992).
 CC -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GnRH family.
 DR PIR: A46030; A46030.
 DR InterPro: IPR002012; GnRH.
 DR Pfam: PF00446; GnRH; 1.

DR PROSITE; PS00473; GNRH; 1.
 KW Hormone; Amidation; Hypothalamus; Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 10 10 AMIDATION.
 SQ SEQUENCE 10 AA; 1204 MW; 284B32337871F5A3 CRC64;

Query Match 25.9%; Score 21; DB 1; Length 10;
 Best Local Similarity 60.0%; Pred. No. 9.7e+02;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Oy 7 WSPVW 11
 :|||
 Db 3 WSHGW 7

RESULT 12

HTF1 PERAM HTF1 PERAM STANDARD; PRT; 8 AA.

AC P04548;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypertrihaloemic factor I (Neuropeptide M-I) (Periplaneta CC-I)
 DE (PEA-CAH-I) (Leb-CC-I) (Hypertrihaloemic neuropeptide I).
 OS Periplaneta americana (American cockroach).
 OS Leptinotarsa decemlineata (Colorado potato beetle), and
 OS Blatta orientalis (Oriental cockroach).
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattodea;
 CC Blattellidae; Periplaneta.
 OX NCBI_TaxID=6978, 7539, 6976;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=P.americana;
 RX MEDLINE=65046530; PubMed=6548628;
 RA Witten J.B., Schaffer M.H., O'Shea X., Cook J.C., Henling M.E.,
 RA Rinehart K.L. Jr.;
 RT "Structures of two cockroach neuropeptides assigned by fast atom
 RT bombardment mass spectrometry."
 RL Biochem. Biophys. Res. Commun. 124:350-358(1984).
 RN [2]
 RP SEQUENCE.
 RC SPECIES=P.americana;
 RX MEDLINE=84298179; PubMed=6591205;
 RA Scarborough R.M., Jamieson G.C., Kalish F., Kramer S.C., McEntee G.A.,
 RA Miller C.A., Schooley E.A.;
 RT "Isolation and primary structure of two peptides with
 RT cardioacceleratory and hyperglycemic activity from the corpora
 RT cardiaca of Periplaneta americana."
 RL Proc. Natl. Acad. Sci. U.S.A. 81:5575-5579(1984).
 RN [3]
 RP SEQUENCE.
 RC SPECIES=B.orientalis; TISSUE=Corpora cardiaca;
 RX MEDLINE=90253659; PubMed=2576128;
 RA Gaede G., Rinehart K.L. Jr.;
 RT "The metabolic neuropeptides of the corpus cardiaca from the potato
 RT beetle and the American cockroach are identical."
 RL Peptides 10:1287-1289(1989).
 RN [4]

RP SEQUENCE.
 RC SPECIES=B.orientalis; TISSUE=Corpora cardiaca;
 RX MEDLINE=90253659; PubMed=2340112;
 RA Gaede G., Rinehart K.L. Jr.;
 RT "Primary structures of hypertrihaloemic neuropeptides isolated from
 RT the corpora cardiaca of the cockroaches Leucophaea maderae,
 RT Gryllodorthina portoricensis, Blattella germanica and Blatta orientalis
 RT and of the stick insect Extatosoma tiaratum assigned by tandem fast
 RT atom bombardment mass spectrometry."
 RL Biol. Chem. Hoppe-Seyler 371:345-354(1990).
 CC -!- FUNCTION: HYPERTRIHALOEMIC FACTORS ARE NEUROPEPTIDES THAT
 CC ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
 CC THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
 CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.

DR PIR; A44960; A44960.
 DR PIR; A49823; A49823.
 DR PIR; S08995; S08995.
 DR InterPro; IPR002047; AKH.
 DR PROSITE; PS00256; AKH; 1.
 KW Neuropeptide; Amidation; Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 8 8 AMIDATION.
 SQ SEQUENCE 8 AA; 991 MW; 86745775B9C452D6 CRC64;

Query Match 24.7%; Score 20; DB 1; Length 8;
 Best Local Similarity 60.0%; Pred. No. 1.3e+05;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 7 WSPVW 11
 :|||
 Db 4 FSPNW 8

RESULT 13

HTF TENMO HTF TENMO STANDARD; PRT; 8 AA.

AC F25419;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE Hypertrihaloemic factor (HOTH) (Hypertrihaloemic neuropeptide).
 DE Tenebrio molitor (Yellow mealworm), and
 OS Zophobas rugipes.
 CC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 CC Neoptera; Endopterygota; Coleoptera; Polyphaga; Cucujiformia;
 CC Tenebrionidae; Tenebrio.
 CX NCBI_TaxID=7067, 7075;
 RN [1]
 RP SEQUENCE.
 RC SPECIES=T.molitor, and Z.rugipes;
 RC TISSUE=Corpora cardiaca;
 RX MEDLINE=90341081; PubMed=2381871;
 RA Gaede G., Rosinski G.;
 RT "The primary structure of the hypertrihaloemic neuropeptide from
 RT tenebrionid beetles: a novel member of the AKH/RPCH family."
 RL Peptides 11:455-459(1990).
 CC -!- FUNCTION: HYPERTRIHALOEMIC FACTORS ARE NEUROPEPTIDES THAT
 CC ELEVATE THE LEVEL OF TREHALOSE IN THE HEMOLYMPH (TREHALOSE IS
 CC THE MAJOR CARBOHYDRATE IN THE HEMOLYMPH OF INSECTS).
 CC -!- SIMILARITY: BELONGS TO THE AKH / HRTH / RPCH FAMILY.

DR PIR; A43976; A43976.
 DR PIR; B43976; B43976.
 DR InterPro; IPR002047; AKH.
 DR PROSITE; PS00256; AKH; 1.
 KW Neuropeptide; Amidation; Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 8 8 AMIDATION.
 SQ SEQUENCE 8 AA; 1005 MW; 86745775B9C44736 CRC64;

Query Match 24.7%; Score 20; DB 1; Length 8;
 Best Local Similarity 60.0%; Pred. No. 1.3e+05;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 7 WSPVW 11
 :|||
 Db 4 FSPNW 8

RESULT 14

RPCH PANBO RPCH PANBO STANDARD; PRT; 8 AA.

AC P08939;
 DT 01-NOV-1988 (Rel. 09, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Red pigment concentrating hormone (RPCH).
 OS Pandanus borealis (Northern red shrimp).

OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
 OC Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea; Pandaloidea;
 OC Pandallidae; Pandalus.
 OX NCBI_TaxID=6703;
 RN 1;
 RP SEQUENCE.
 RX MEDLINE=75054965; PubMed=4433569;
 RA Fernlund P.;
 RT "Structure of the red-pigment-concentrating hormone of the shrimp,
 RL Pandalus borealis.";
 RI Biochim. Biophys. Acta 371:304-311(1974).
 CC -!- FUNCTION: THIS HORMONE ADAPTS THE ANIMAL TO LIGHT BACKGROUNDS BY
 CC STIMULATING CONCENTRATION OF THE PIGMENT OF ITS RED BODY-
 CC CHROMATOPHORES.
 CC -!- SIMILARITY: BELONGS TO THE AXH / HRTH / RPCH FAMILY.
 DR PIR; A61348; A61348.
 DR InterPro: IPR002047; AKH.
 DR PROSITE: PS00256; AKH; 1.
 KW Pigment; Hormone; Amidation; Pyrrolidone carboxylic acid.
 FT MOD_RES 1 2
 FT MOD_RES 8 8
 SQ SEQUENCE 8 AA; 948 MW; 86786775B9C44736 CRC64;

 Query Match 24.7%; Score 20; DB 1; Length 8;
 Best Local Similarity 60.0%; Pred. No. 1.3e-05;
 Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

 QY 7 WSPVW 11
 Db :|||
 Db 4 FSPGW 8

 RESULT 15
 GON3_ONCKE STANDARD; PR7; 10 AA.
 AC P20367; P81751;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Gonadoliberin III (Gonadotropin-releasing hormone III) (GNRH-III) (LH-
 DE RH III) (Luliberin III).
 GN GNRH3.
 OS Oncorhynchus keta (Chum salmon), and
 OS Clupea pallasi (Pacific herring)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
 OX NCBI_TaxID=8018, 30724;
 RN 1;
 RN SEQUENCE.
 RC SPECIES=C.keta;
 RX MEDLINE=83195140; PubMed=6341999;
 RA Sherwood N., Eiden L., Brownstein M., Spiess J., Rivier J., Vale W.;
 RT "Characterization of a teleost gonadotropin-releasing hormone.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:2794-2798(1983).
 RN 1;
 RN SEQUENCE, AND FUNCTION.
 RC SPECIES=C.pallasi; TISSUE=Brain, and Pituitary;
 RX MEDLINE=20114351; PubMed=1650929;
 RA Sherwood J., Powell J.F.P., Park M., Fischer W.H., Craig A.G.,
 RA Chang C.P., Rivier J.E., Sherwood N.M.;
 RT "Primary structure and function of three gonadotropin-releasing
 hormones, including a novel form, from an ancient teleost, herring.";
 RL Endocrinology 141:505-512(2000).
 CC -!- FUNCTION: Stimulates the secretion of gonadotropins; it stimulates
 CC the secretion of both luteinizing and follicle-stimulating
 CC hormones.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the GNRH family.
 DR PIR; A21114; A21114.
 DR InterPro: IPR002012; GNRH.
 DR Pfam; PF00446; GNRH; 1.
 DR PROSITE; PS00473; GNRH; 1.

KW Hormone; Amidation; Hypothalamus; Pyrrolidone carboxylic acid.
 FT MOD_RES 1 1
 FT MOD_RES 10 10
 FT MOD_RES 10 10
 SQ SEQUENCE 10 AA; 1230 MW; 284B3233786B45A3 CRC64;

 Query Match 24.7%; Score 20; DB 1; Length 10;
 Best Local Similarity 60.0%; Pred. No. 1.3e-03;
 Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

 QY 7 WSPVW 11
 Db :|||
 Db 3 WSYGW 7

 Search completed: October 22, 2003, 20:52:27
 Job time : 23 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 20:49:21 ; Search time 95 Seconds
(without alignments)
32.596 Million cell updates/sec

Title: US-09-995-804A-2

Perfect score: 81

Sequence: 1 THSPPMKSPWP 12

Scoring table: ELOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 2287

Minimum DB seq length: 0

Maximum DB seq length: 12

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL23:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_thc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_redent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_rvirus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	38	46.9	9	13	Q9PRJ4 lepisosteus
2	32	39.5	10	8	Q8SHB1
3	32	39.5	10	8	Q8SH99
4	29	35.8	10	8	Q8SHL3
5	29	35.8	10	8	Q8SHK1
6	29	35.8	10	8	Q8SHP0
7	29	35.8	10	8	Q8SHM2
8	29	35.8	10	8	Q8SHL0
9	27	33.3	11	7	Q7R898
10	27	33.3	11	7	Q7R894
11	25	30.9	11	7	Q7R896
12	25	30.9	11	8	Q8VAZ1
13	25	30.9	11	8	Q8VB39
14	25	30.9	11	8	Q8MB58
15	25	30.9	11	8	Q8MAZ3
16	25	30.9	11	8	Q8MBE1

17	25	30.9	11	8	Q8MB77	Q8mb77 odonellia b
18	25	30.9	11	8	Q8MB79	Q8mb79 aniseia arg
19	25	30.9	11	8	Q8MB97	Q8mb97 merretia pe
20	24	29.6	10	8	Q8SHC6	Q8shc6 furcifer be
21	24	29.6	11	7	Q7R895	Q77895 oreochromis
22	23	28.4	8	11	Q62721	Q62721 rattus norv
23	23	28.4	10	8	Q8SHF6	Q8shf6 chamaeleo m
24	22	27.2	8	11	Q9ET18	Q9et18 mus spratus
25	22	27.2	8	11	Q9ET17	Q9et17 mus caroli
26	22	27.2	8	11	Q9ET16	Q9et16 mesocricetu
27	22	27.2	8	12	Q64971	Q64971 alfalfa mos
28	22	27.2	10	8	Q8SHB4	Q8shb4 furcifer la
29	22	27.2	10	8	Q8SHC3	Q8shc3 furcifer la
30	22	27.2	10	8	Q8SHC0	Q8shc0 furcifer la
31	22	27.2	10	8	Q8SHB7	Q8shb7 furcifer ou
32	21	25.9	10	2	Q8KJF1	Q8rjf1 pseudomonas
33	21	25.9	10	2	Q9R5T2	Q9rst2 acetobacter
34	21	25.9	10	8	Q8SHG2	Q8shg2 chamaeleo j
35	21	25.9	10	8	Q8SHN1	Q8shn1 bradypodion
36	21	25.9	12	6	P81127	P81127 bos indicus
37	20	24.7	9	4	Q9H326	Q9h326 homo sapien
38	20	24.7	11	4	Q9UCR1	Q9ucr1 homo sapien
39	20	24.7	11	11	Q61797	Q61797 mus musculu
40	20	24.7	12	4	Q8IVH0	Q8ivh0 homo sapien
41	20	24.7	12	7	Q7R915	Q77915 oreochromis
42	20	24.7	12	10	Q93X21	Q93x21 zea mays fm
43	20	24.7	12	13	Q9PSW5	Q9psw5 gallus gail
44	19	23.5	8	4	Q8IVK3	Q8ivk3 homo sapien
45	19	23.5	9	2	Q9R635	Q9r635 cnamylidia t

ALIGNMENTS

RESULT 1

Q9PRJ4 PRELIMINARY; PRT; 9 AA.
 AC Q9PRJ4;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE Bradykinin.
 OS Lepisosteus osseus (Long-nosed gar), and
 OS Aria calva (Bowfin).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Semionotiformes; Lepisosteidae;
 OC Lepisosteus.
 OX NCBI_TaxID=34771, 7924;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=95380361; PubMed=7651903;
 RA Conlon J.M., Platzack B., Marra J.E., Youson J.H., Olson K.R.;
 RT "Isolation and biological activity of [trp5]bradykinin from the plasma
 of the phylogenetically ancient fish, the bowfin and the longnosed
 gar.";
 RL Peptides 16:485-489(1995).
 SQ SEQUENCE 9 AA; 1099 MW; 3393D775A3786777 CRC64;

Query Match 46.9%; Score 38; DB 13; Length 9;
 Best Local Similarity 85.7%; Pred. NO. 8.3e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 3 RPPMKSP 9
 |||||
 Db 1 RPPGWSP 7

RESULT 2

Q8SHB1 PRELIMINARY; PRT; 10 AA.
 ID Q8SHB1
 AC Q8SHB1;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)

```

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update);
DE Cytochrome c oxidase subunit I (fragment);
GN COI.
OS Rhampholeon brevicaudatus.
CG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae;
OC Rhampholeon.
OX NCBI_TaxID=9912;
RN [1]
RP SEQUENCE FROM N.A.
RA Townsend T.M., Larson A.L.;
RT "Molecular Phylogenetics and Mitochondrial Genomic Evolution in the
RT Chamaeleonidae (Reptilia, Squamata).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF448771; AAL90595.1; -.
KW Mitochondrion.
FT NON-TER
SQ SEQUENCE 10 AA; 1291 MW; 86218E2733641771 CRC64;

Query Match 39.5%; Score 32; DB 9; Length 10;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 MWSPVW 11
DB 1 MWTPRW 6

RESULT 3
Q8SH99 PRELIMINARY; PRT; 10 AA.
AC Q8SH99;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cytochrome c oxidase subunit I (fragment).
GN COI.
OS Brookesia nasus.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae; Brookesia.
CX NCBI_TaxID=179893;
RN [1]
RP SEQUENCE FROM N.A.
RA Townsend T.M., Larson A.L.;
RT "Molecular Phylogenetics and Mitochondrial Genomic Evolution in the
RT Chamaeleonidae (Reptilia, Squamata).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF448775; AAL90607.1; -.
KW Mitochondrion.
FT NON-TER
SQ SEQUENCE 10 AA; 1261 MW; C7218E273364176D CRC64;

Query Match 39.5%; Score 32; DB 9; Length 10;
Best Local Similarity 66.7%; Pred. No. 1.5e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 6 MWSPVW 11
DB 1 MWTPRW 6

RESULT 4
Q8SHL3 PRELIMINARY; PRT; 10 AA.
AC Q8SHL3;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cytochrome c oxidase subunit I (fragment).
GN COI.
OS Calumma farcifer.

```

```

CG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae; Calumma.
OX NCBI_TaxID=179902;
RN [1]
RP SEQUENCE FROM N.A.
RA Townsend T.M., Larson A.L.;
RT "Molecular Phylogenetics and Mitochondrial Genomic Evolution in the
RT Chamaeleonidae (Reptilia, Squamata).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF448736; AAL90490.1; -.
KW Mitochondrion.
FT NON-TER
SQ SEQUENCE 10 AA; 1303 MW; 5E218E2733641767 CRC64;

Query Match 35.8%; Score 29; DB 8; Length 10;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 MWSPVW 11
DB 1 MWTPRW 6

RESULT 5
Q8SHK1 PRELIMINARY; PRT; 10 AA.
AC Q8SHK1;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cytochrome c oxidase subunit I (fragment).
GN COI.
OS Calumma nasuta.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae; Calumma.
OX NCBI_TaxID=179905;
RN [1]
RP SEQUENCE FROM N.A.
RA Townsend T.M., Larson A.L.;
RT "Molecular Phylogenetics and Mitochondrial Genomic Evolution in the
RT Chamaeleonidae (Reptilia, Squamata).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF448740; AAL90502.1; -.
KW Mitochondrion.
FT NON-TER
SQ SEQUENCE 10 AA; 1303 MW; 5E218E2733641767 CRC64;

Query Match 35.8%; Score 29; DB 8; Length 10;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 MWSPVW 11
DB 1 MWTPRW 6

RESULT 6
Q8SHP0 PRELIMINARY; PRT; 10 AA.
AC Q8SHP0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Cytochrome c oxidase subunit I (fragment).
GN COI.
OS Bradypodion adolfi-friederici.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae;
OC Bradypodion.
OX NCBI_TaxID=179885;

```

```

RN [1].
RP SEQUENCE FROM N.A.
RA Townsend T.M., Larson A.L.;
RT "Molecular Phylogenetics and Mitochondrial Genomic Evolution in the
RT Chamaeleonidae (Reptilia, Squamata).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF448727; AAL90463.1; -.
KW Mitochondrion.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1303 MW; 5E218E2733641767 CRC64;

Query Match 35.8%; Score 29; DB 8; Length 10;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 WSPVW 11
DB 1 MWLPRW 6

RESULT 7
Q8SHM2 PRELIMINARY; PRT; 10 AA.
AC Q8SHM2;
DT 01-JUN-2002 (TrEMBLrel. 21, Created);
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update);
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update);
DE Cytochrome c oxidase subunit I (fragment).
GN COI.
OS Calumma boettgeri.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae; Calumma.
OX NCBI_TaxID=179899;
RN [1]
RP SEQUENCE FROM N.A.
RA Townsend T.M., Larson A.L.;
RT "Molecular Phylogenetics and Mitochondrial Genomic Evolution in the
RT Chamaeleonidae (Reptilia, Squamata).";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF448733; AAL90481.1; -.
KW Mitochondrion.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1303 MW; 5E218E2733641767 CRC64;

Query Match 35.8%; Score 29; DB 8; Length 10;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 WSPVW 11
DB 1 MWLPRW 6

RESULT 8
Q8SHLO PRELIMINARY; PRT; 10 AA.
AC Q8SHLO;
DT 01-JUN-2002 (TrEMBLrel. 21, Created);
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update);
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update);
DE Cytochrome c oxidase subunit I (fragment).
GN COI.
OS Calumma gastrotaenia.
OG Mitochondrion.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Iguania; Acrodonta; Chamaeleonidae; Calumma.
OX NCBI_TaxID=179903;
RN [1]
RP SEQUENCE FROM N.A.
RA Townsend T.M., Larson A.L.;
RT "Molecular Phylogenetics and Mitochondrial Genomic Evolution in the
RT Chamaeleonidae (Reptilia, Squamata).";

```

```

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF448737; AAL90493.1; -.
KW Mitochondrion.
FT NON_TER 10
SQ SEQUENCE 10 AA; 1303 MW; 5E218E2733641767 CRC64;

Query Match 35.8%; Score 29; DB 8; Length 10;
Best Local Similarity 66.7%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 6 WSPVW 11
DB 1 MWLPRW 6

RESULT 9
O77898 PRELIMINARY; PRT; 11 AA.
AC O77898;
DT 01-NOV-1998 (TrEMBLrel. 08, Created);
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update);
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update);
DE MHC class II B locus 12 (fragment).
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Malaga-Trillo E., Zaleska-Ruczyńska Z., McAndrew B., Vincek V.,
RA Figueroa F., Sultmann H., Klein J.;
RT "Linkage relationships and haplotype polymorphism among cichlid mhc
RT class II B loci.";
RL Genetics 149:1527-1537(1998).
DR EMBL: AF050008; AAC41347.1; -.
FT NON_TER 11
SQ SEQUENCE 11 AA; 1367 MW; 3F47D0A62C045A3 CRC64;

Query Match 33.3%; Score 27; DB 7; Length 11;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 WSPVW 11
DB 4 WSIWV 8

RESULT 10
O77894 PRELIMINARY; PRT; 11 AA.
AC O77894;
DT 01-NOV-1998 (TrEMBLrel. 08, Created);
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update);
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update);
DE MHC class II B locus 12 (fragment).
OS Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
OC Cichlidae; Oreochromis.
OX NCBI_TaxID=8128;
RN [1]
RP SEQUENCE FROM N.A.
RA Malaga-Trillo E., Zaleska-Ruczyńska Z., McAndrew B., Vincek V.,
RA Figueroa F., Sultmann H., Klein J.;
RT "Linkage relationships and haplotype polymorphism among cichlid mhc
RT class II B loci.";
RL Genetics 149:1527-1537(1998).

```

```

DR  EXBL: AF050004; AAC43343.1; -.
FT  NON_TER 11
SC  SEQUENCE 11 AA; 1367 MW; 3F47DC0A62C045A3 CRC64;

Query Match      33.3%; Score 27; DB 7; Length 11;
Best Local Similarity 80.0%; Pred. No. 8.9e+02;
Matches 4; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY  7 WSPVW 11
DB  4 WSIW 8

RESULT 11
O77896
ID  O77896 PRELIMINARY; PRT; 11 AA.
AC  O77896;
DT  01-NOV-1998 (TREMBlrel. 08, Created)
DT  01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT  01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE  MHC class II B locus 12 (Fragment).
OS  Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoeleostei;
OC  Acanthomorpha; Acanthopterygii; Perciformes; Labroidae;
OC  Cichlidae; Oreochromis.
OX  NCBI_TaxID=8128;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=98315113; PubMed=9649539;
RA  Malaga-Trillo E., Zaleska-Rutczynska Z., McAndrew B., Vincek V.,
RA  Figueroa F., Sultmann H., Klein J.,
RT  "Linkage relationships and haplotype polymorphism among cichlid mhc
RT  class II B loci."
RL  Genetics 149:1527-1537(1998).
DR  EMBL; AF050006; AAC43345.1; -.
FT  NON_TER 11
FT  NON_TER 1
SQ  SEQUENCE 11 AA; 1399 MW; 3F47DB7A72665A3 CRC64;

Query Match      30.9%; Score 25; DB 7; Length 11;
Best Local Similarity 60.0%; Pred. No. 1.7e+03;
Matches 3; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY  7 WSPVW 11
DB  4 WSIW 8

RESULT 12
OBMAZ1
ID  OBMAZ1 PRELIMINARY; PRT; 11 AA.
AC  OBMAZ1;
DT  01-OCT-2002 (TREMBlrel. 22, Created)
DT  01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT  01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE  PsbJ (Fragment).
GN  PsbJ.
OS  Maripa paniculata.
OC  Chloroplast.
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC  Asteridae; lamids; Solanales; Convolvulaceae; Maripa.
OX  NCBI_TaxID=197411;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Stefanovic S., Krueger L., Olmstead R.G.;
RT  "Monophyly of the Convolvulaceae and circumscription of their major
RT  lineages based on DNA sequences of multiple chloroplast loci."
RL  Am. J. Bot. 0:0-0(2002).
DR  EMBL; AY100937; AAY55869.1; -.
KW  Chloroplast.

```

```

FT  NON_TER 11
SC  SEQUENCE 11 AA; 1260 MW; 93736D59440861B1 CRC64;

Query Match      30.9%; Score 25; DB 8; Length 11;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY  1 THRPMMW 7
DB  5 TGRIPLM 11

RESULT 13
Q8MB39
ID  Q8MB39 PRELIMINARY; PRT; 11 AA.
AC  Q8MB39;
DT  01-OCT-2002 (TREMBlrel. 22, Created)
DT  01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT  01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE  PsbJ (Fragment).
GN  PsbJ.
OS  Wilsonia humilis.
OC  Chloroplast.
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC  Asteridae; lamids; Solanales; Convolvulaceae; Wilsonia.
OX  NCBI_TaxID=197481;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Stefanovic S., Krueger L., Olmstead R.G.;
RT  "Monophyly of the Convolvulaceae and circumscription of their major
RT  lineages based on DNA sequences of multiple chloroplast loci."
RL  Am. J. Bot. 0:0-0(2002).
DR  EMBL; AY100914; AAM55777.1; -.
FT  NON_TER 11
FT  NON_TER 1
SQ  SEQUENCE 11 AA; 1260 MW; 93736D59440861B1 CRC64;

Query Match      30.9%; Score 25; DB 8; Length 11;
Best Local Similarity 57.1%; Pred. No. 1.7e+03;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY  1 THRPMMW 7
DB  5 TGRIPLM 11

RESULT 14
Q8MB58
ID  Q8MB58 PRELIMINARY; PRT; 11 AA.
AC  Q8MB58;
DT  01-OCT-2002 (TREMBlrel. 22, Created)
DT  01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT  01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE  PsbJ (Fragment).
GN  PsbJ.
OS  Seddera hirsuta.
OC  Chloroplast.
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC  Asteridae; lamids; Solanales; Convolvulaceae; Seddera.
OX  NCBI_TaxID=197444;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Stefanovic S., Krueger L., Olmstead R.G.;
RT  "Monophyly of the Convolvulaceae and circumscription of their major
RT  lineages based on DNA sequences of multiple chloroplast loci."
RL  Am. J. Bot. 0:0-0(2002).
DR  EMBL; AY100905; AAM55743.1; -.
KW  Chloroplast.
FT  NON_TER 11
FT  NON_TER 1
SQ  SEQUENCE 11 AA; 1260 MW; 93736D59440861B1 CRC64;

```

Query Match 30.9%; Score 25; DB 8; Length 11;
 Best Local Similarity 57.1%; Pred. No. 1.7e+03;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 THRPMM 7
 | | | | |
 Db 5 TGRPLW 11

RESULT 15

Q8MAZ3
 ID Q8MAZ3 PRELIMINARY; PRT; 11 AA.
 AC Q8MAZ3;
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE PSBJ (Fragment).
 GN PSBJ.
 OS Maripa repens.
 OG Chloroplast.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 OC Asteridae; -amids; Solanales; Convolvulaceae; Maripa.
 OX NCBI_TaxID=197412;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Stefanovic S., Krueger L., Olmstead R.G.;
 RT "Monophyly of the Convolvulaceae and circumscription of their major
 R: lineages based on DNA sequences of multiple chloroplast loci,"
 RL Am. J. Bot. 0:0-0(2002).
 DR EMBL; AY100936; AAM55865.1; -.
 KW Chloroplast.
 FT NON_TER
 SQ SEQUENCE 11 AA; 1260 MW; 9373659440861B1 CRC64;

Query Match 30.9%; Score 25; DB 8; Length 11;
 Best Local Similarity 57.1%; Pred. No. 1.7e+03;
 Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 THRPMM 7
 | | | | |
 Db 5 TGRPLW 11

Search completed: October 22, 2003, 20:54:11
 Job time : 99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 20:41:56 ; Search time 82 Seconds
(without alignments)
23,228 Million cell updates/sec

Title: US-09-995-804A-2
Perfect score: 81
Sequence: 1 THRPWKSVPWP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 287032

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Vatch 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	81	100.0	12	23	AAU98658 Peptide #2 capable
2	41	50.6	12	21	AAU98819 Rhodamine binding
3	41	50.6	12	21	AAU988175 Rhodamine binding
4	41	50.6	12	21	AAU988199 Rhodamine binding
5	38	46.9	9	21	AAU989374 Bradykinin peptide
6	35	43.2	8	14	AAU935961 Hepatitis C virus
7	35	43.2	8	22	AAU901137 Hepatitis C virus
8	35	43.2	8	22	AAU901682 Hepatitis C virus
9	35	43.2	9	22	AAU901136 Hepatitis C virus

12	35	43.2	9	22	AAU901801 Hepatitis C virus
11	35	43.2	10	18	AAU94284 CBF1/RBPJk interac
13	35	43.2	10	22	AAU901093 Hepatitis C virus
14	35	43.2	10	22	AAU901138 Hepatitis C virus
15	35	43.2	10	22	AAU901848 Hepatitis C virus
16	35	43.2	11	22	AAU901581 Hepatitis C virus
17	35	43.2	11	22	AAU902000 Hepatitis C virus
18	35	43.2	12	22	AAU90030 Internalising pept
19	34	42.0	12	21	AAU91767 I2-1 antagonist pe
20	33	40.7	12	18	AAU94283 CBF1/RBPJk interac
21	33	40.7	12	18	AAU916063 Peptide containing
22	33	40.7	12	19	AAU98701 Peptide binding in
23	33	40.7	12	19	AAU98180 IL-132 binding pe
24	33	40.7	12	20	AAU99826 Interleukin-1 type
25	32	39.5	12	21	AAU94983 Amino acid sequenc
26	32	39.5	7	22	AAU946075 Human TF anti-Idio
27	32	39.5	9	20	AAU989146 ScPV (dig) variant
28	32	39.5	10	20	AAU948151 Immunogenic peptid
29	31	38.3	12	13	AAU94222 Partition coeffici
30	31	38.3	9	20	AAU948126 Immunogenic peptid
31	31	38.3	10	11	AAU90505 Bradykinin antagon
32	31	38.3	10	17	AAU916991 ABL SH3 domain-bin
33	31	38.3	10	18	AAU925352 ABL SH3 domain bin
34	31	38.3	10	19	AAU950231 Tumour-related sug
35	31	38.3	10	24	ABJ22105 162PIE6 cancer gen
36	31	38.3	10	24	ABJ22806 162PIE6 cancer gen
37	31	38.3	10	24	ABJ24870 162PIE6 cancer gen
38	31	38.3	11	19	AAU918162 Bi-BK receptor ant
39	31	38.3	12	23	AAU93484 Bi-BK receptor ant
40	30.5	37.7	12	15	AAU960455 Antiproliferative
41	30.5	37.7	12	15	AAU960440 Antiproliferative
42	30	37.0	7	18	AAU95025 Immunomodulatory p
43	30	37.0	7	19	AAU951154 Cysteine containin
44	30	37.0	7	20	AAU909439 Immunoactive pepti
45	30	37.0	9	16	AAU97293 Hepatitis C virus

ALIGNMENTS

RESULT 1

AAU98658
ID AAU98658 standard; peptide; 12 AA.

XX AAU98658;

XX 24-SEP-2002 (first entry)

XX Peptide #2 capable of binding to and internalising htrR.

XX Human transferrin receptor; htrR; malignancy; oral cancer;

XX liver cancer; pancreatic cancer; prostate cancer; cytosatic.

XX Synthetic.

XX WO200244329-A2.

XX 06-JUN-2002.

XX 29-NOV-2001; 2001WO-US44518.

XX 30-NOV-2000; 2000US-253940P.

XX (UAB) UAB RES FOUND.

XX Engler J, Lee JH, Collawan J, Moore B;

XX WPI; 2002-508504/54.

XX Novel peptide useful for targeting other peptides and proteins into cells expressing human transferrin receptor

self pd

```

PS Claim 1; Page 19; 26pp; English.
XX
CC The present invention relates to peptides capable of binding to and
CC internalising with the human transferrin receptor (hTfR). The peptides
CC of the invention are useful for targeting other peptides and proteins
CC into cells expressing human transferrin receptor (hTfR), and also for
CC facilitating entry of diagnostically and therapeutically useful agents,
CC including peptides and proteins. They are particularly useful for the
CC study and treatment of malignancies in cancers (e.g. oral, liver,
CC pancreatic, and prostate cancers). The present sequence represents a
CC peptide that is capable of binding to and internalising hTfR.
XX
SQ Sequence 12 AA;
XX
XX Query Match 100.0%; Score 81; DB 23; Length 12;
XX Best Local Similarity 100.0%; Pred. No. 4.4e-35;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 THRPPMSPWP 12
DB 1 THRPPMSPWP 12
XX
RESULT 2
ID AAY88119 standard; peptide; 12 AA.
XX
AC AAY88119;
XX
DT 17-OCT-2000 (first entry)
XX
DE Rhodamine binding peptide fluorettre #2.
XX
KW Fluorettre; fluorophore dye; fluorescent; detection; diagnostic.
XX
OS Synthetic.
XX
PN WC200023463-A2.
XX
PD 27-APR-2000.
XX
PF 15-OCT-1999; 99WO-US24266.
XX
PR 16-OCT-1998; 98US-0104465.
XX
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI Nolan GP, Rozinov MN;
XX
WP1; 2000-339647/29.
XX
XX Novel fluorettre peptides which bind with high specificity to
XX fluorophore dyes, useful for detecting biological materials, molecules,
XX target analytes, and intra- and inter-molecular interactions -
XX
PS Claim 9; Page 19; 64pp; English.
XX
CC This invention describes a novel peptide (I) termed a fluorettre which
CC comprises a sequence of at least 8 amino acids and which binds to a
CC fluorophore dye. The fluorettres maybe used in detecting biological
CC materials, molecules or target analytes, intracellular events, and intra-
CC and inter-molecular interactions, as well as discovering effective
CC inhibitors. They may also be used for in vitro assays using a large set
CC of fluorophore dyes, e.g. in detection of viral and delivery systems,
CC diagnostics and high throughput assays. The new peptides, due to their
CC small size, are not intrusive to the systems being studied hence, permit
CC detection and analysis of a target moiety or molecule while minimizing
CC target modification. Unlike conventional techniques such as the
CC light-emitting technique, the use of fluorettres minimize or avoid
CC substantial genetic manipulations which may disrupt, interfere, or alter
CC the process being measured. This sequence represents a peptide fluorettre
CC capable of binding to Rhodamine which is described in the method of the
CC invention.
XX
SQ Sequence 12 AA;
XX
XX Query Match 100.0%; Score 81; DB 23; Length 12;
XX Best Local Similarity 100.0%; Pred. No. 4.4e-35;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 THRPPMSPWP 12
DB 1 THRPPMSPWP 12
XX
RESULT 3
ID AAY88175 standard; peptide; 12 AA.
XX
AC AAY88175;
XX
DT 17-OCT-2000 (first entry)
XX
DE Rhodamine binding peptide fluorettre from phage clone RhRS308.
XX
KW Fluorettre; fluorophore dye; fluorescent; detection; diagnostic.
XX
OS Synthetic.
XX
PN WC200023463-A2.
XX
PD 27-APR-2000.
XX
PF 15-OCT-1999; 99WO-US24266.
XX
PR 16-OCT-1998; 98US-0104465.
XX
PA (STRD ) UNIV LELAND STANFORD JUNIOR.
XX
PI Nolan GP, Rozinov MN;
XX
WP1; 2000-339647/29.
XX
XX Novel fluorettre peptides which bind with high specificity to
XX fluorophore dyes, useful for detecting biological materials, molecules,
XX target analytes, and intra- and inter-molecular interactions -
XX
PS Example 3; Page 41; 64pp; English.
XX
CC This invention describes a novel peptide (I) termed a fluorettre which
CC comprises a sequence of at least 8 amino acids and which binds to a
CC fluorophore dye. The fluorettres maybe used in detecting biological
CC materials, molecules or target analytes, intracellular events, and intra-
CC and inter-molecular interactions, as well as discovering effective
CC inhibitors. They may also be used for in vitro assays using a large set
CC of fluorophore dyes, e.g. in detection of viral and delivery systems,
CC diagnostics and high throughput assays. The new peptides, due to their
CC small size, are not intrusive to the systems being studied hence, permit
CC detection and analysis of a target moiety or molecule while minimizing
CC target modification. Unlike conventional techniques such as the
CC light-emitting technique, the use of fluorettres minimize or avoid
CC substantial genetic manipulations which may disrupt, interfere, or alter
CC the process being measured. This sequence represents a peptide fluorettre
CC capable of binding to Rhodamine which is described in the method of the
CC invention.
XX
SQ Sequence 12 AA;
XX
XX Query Match 100.0%; Score 81; DB 23; Length 12;
XX Best Local Similarity 100.0%; Pred. No. 4.4e-35;
XX Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 2 HRPMPWSPWP 11
DB 3 HRPMPWTPVF 12
XX

```

RESULT 4
 AAY88199
 ID AAY88199 standard; peptide; 12 AA.
 XX
 AC AAY88199;
 DT 17-OCT-2000 (first entry)
 DE Rhodamine binding peptide fluorette from phage clone RARS308.
 XX
 KW Fluorette; fluorophore dye; fluorescent; detection; diagnostic.
 XX
 CS Synthetic.
 PN W0200023463-A2.
 XX
 PD 27-APR-2000.
 XX
 PF 15-OCT-1999; 99WO-US24266.
 XX
 PR 16-OCT-1998; 98US-0104465.
 XX
 PA (STRD) UNIV LELAND STANFORD JUNIOR.
 XX
 PI Nolan GP, Rozinov MN;
 XX
 PZ
 DR WP1; 2000-339647/29.
 XX
 PT Novel fluorette peptides which bind with high specificity to
 PT fluorophore dyes, useful for detecting biological materials, molecules,
 PT target analytes, and intra- and inter-molecular interactions -
 XX
 PS Example 4; Page 45; 64pp; English.
 XX
 CC This invention describes a novel peptide (1) termed a fluorette which
 CC comprises a sequence of at least 8 amino acids and which binds to a
 CC fluorophore dye. The fluorettes may be used in detecting biological
 CC materials, molecules or target analytes, intracellular events, and intra-
 CC and inter-molecular interactions, as well as discovering effective
 CC inhibitors. They may also be used for in vitro assays using a large set
 CC of fluorophore dyes, e.g. in detection of viral and delivery systems,
 CC diagnostics and high throughput assays. The new peptides, due to their
 CC small size, are not intrusive to the systems being studied hence, permit
 CC detection and analysis of a target moiety or molecule while minimizing
 CC target modification. Unlike conventional techniques such as the
 CC light-emitting technique, the use of fluorettes minimize or avoid
 CC substantial genetic manipulations which may disrupt, interfere, or alter
 CC the process being measured. This sequence represents a peptide fluorette
 CC capable of binding to Rhodamine which is described in the method of the
 CC invention.
 XX
 SQ Sequence 12 AA;
 Query Match 50.6%; Score 41; DB 21; Length 12;
 Best Local Similarity 60.0%; Pred. No. 16;
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
 QY 2 HRPMPGSPWV 11
 DB 3 HRPMTWTPVF 12
 XX
 RESULT 5
 AAY59374
 ID AAY59374 standard; peptide; 9 AA.
 XX
 AC AAY59374;
 XX
 DT 13-MAR-2000 (first entry)
 DE Bradykinin peptide analogue Brdy1.
 XX

KW Neurotensin analogue; neo-tryptophan; neurotensin response;
 KW antinociception; hypothermia; appetite reduction; body weight reduction;
 KW weight gain reduction; catalepsy; central nervous system; schizophrenia;
 KW antipsychotic effect; CNS stimulant; therapy; bradykinin.
 XX
 CS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 5 /note= "neo-Trp"
 FT
 XX
 PN W09952539-A1.
 XX
 PD 21-OCT-1999.
 XX
 PF 09-APR-1999; 99WO-US07810.
 XX
 PR 10-APR-1998; 98US-0081356.
 PR 09-JUL-1998; 98US-0092195.
 PR 27-AUG-1998; 98US-0098119.
 PR 14-DEC-1998; 98US-0112137.
 XX
 PA (MAYO-) MAYO FOUNDATION.
 XX
 PI Richelson E, Cusack BM, Pang Y, McCormick DJ, Faug A, Tyler BM;
 PZ Boules M;
 XX
 DR WP1; 2000-061969/05.
 XX
 PT Neo-tryptophan, its derivatives and polypeptides containing
 PT neo-tryptophan -
 XX
 PS Claim 7; Page 19; 65pp; English.
 XX
 CC This sequence represents a angiotensin analogue. The invention
 CC relates to the new amino acid neo-tryptophan. Polypeptides containing
 CC neo-tryptophan are useful for inducing a neurotensin response in mammals,
 CC particularly humans. The response comprises antinociception, hypothermia,
 CC a reduction in appetite, a reduction in body weight, a reduction in body
 CC weight gain, prevention or reduction of catalepsy, reduction of an effect
 CC of a central nervous system (CNS) stimulant, an antipsychotic effect
 CC (particularly reduction of the signs or symptoms of schizophrenia) or
 CC interaction with a neurotensin receptor (particularly rat neurotensin
 CC receptor or human neurotensin receptor).
 XX
 SQ Sequence 9 AA;
 Query Match 46.9%; Score 38; DB 21; Length 9;
 Best Local Similarity 85.7%; Pred. No. 9.3e+05;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 RPPMWSP 9
 DB 1 RPPGWSP 7
 XX
 RESULT 6
 AAR35961
 ID AAR35961 standard; protein; 8 AA.
 XX
 AC AAR35961;
 XX
 DT 25-MAR-2003 (updated)
 DT 24-MAY-1993 (first entry)
 XX
 DE Hepatitis C virus (HCV) epitope Ep26.
 XX
 KW Hepatitis; liver disease; HCV; monoclonal antibody; epitope;
 KW immobilised reagent; immunoassay; diagnosis; detection; treatment;
 KW infection.
 XX
 OS Hepatitis C virus type 1.
 XX

PV WO9300365-A2.
XX 07-JAN-1993.
XX 24-JUN-1992; 52WO-JS05388.
XX 24-JUN-1991; 91US-0722489.
XX (CHIR) CHIRON CORP.
XX Chien DY, Rutter W;
XX WPI; 1993-036334/34.
XX Polypeptide(s) comprising truncated hepatitis C virus sequences -
PT for detection, prevention and treatment of hepatitis C infection
XX
XX Example A; Page 36; 80pp; English.
XX This octamer was found to be immunoreactive with anti-HCV anti-sera.
CC In the epitope mapping experiment three different samples of anti-sera
CC were reacted with the peptide octamer, and then incubated with
CC HRP-labelled goat anti-human Ig anti-sera, to enable detection of
CC binding. This epitope starts from amino acid 1601 of the HCV
CC polyprotein.
CC This was found to be a strong epitope.
CC (Updated on 25-MAR-2003 to correct PN field.)
XX
XX Sequence 8 AA;
SQ
Query Match 43.2%; Score 35; DB 14; Length 8;
Best Local Similarity 50.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 4 PPMWSPVW 11
DB 1 PPSWDQMW 8
RESULT 7
AAJ01137
ID AAJ01137 standard; Peptide; 8 AA.
XX AC
XX AAJ01137;
D7 02-JUL-2001 (first entry)
XX Hepatitis C virus epitope #1128.
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX Hepatitis C virus.
OS
XX W02C0121:89-Al.
PN 29-MAR-2001.
XX 19-JUL-2000; 20CWO-US19774.
XX 19-JUL-1999; 99US-0357737.
XX (EPIM-) EPIMMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX A new composition useful as a vaccine against hepatitis C virus -
PS Disclosure; Page 128; 214pp; English.
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention.
XX
XX Sequence 8 AA;
SQ
Query Match 43.2%; Score 35; DB 14; Length 8;
Best Local Similarity 50.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 4 PPMWSPVW 11
DB 1 PPSWDQMW 8
RESULT 8
AAJ01682
ID AAJ01682 standard; Peptide; 8 AA.
XX AC
XX AAJ01682;
XX 02-JUL-2001 (first entry)
XX Hepatitis C virus epitope #1673.
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX Hepatitis C virus.
OS
XX W02C0121:189-Al.
PN 29-MAR-2001.
XX 19-JUL-2000; 20CWO-US19774.
XX 19-JUL-1999; 99US-0357737.
XX (EPIM-) EPIMMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX A new composition useful as a vaccine against hepatitis C virus -
PS Disclosure; Page 128; 214pp; English.
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention.
XX
XX Sequence 8 AA;
SQ
Query Match 43.2%; Score 35; DB 22; Length 8;
Best Local Similarity 50.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 4 PPMWSPVW 11
DB 1 PPSWDQMW 8
RESULT 9
AAJ01136
ID AAJ01136 standard; Peptide; 9 AA.
XX AC
XX AAJ01136;

CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention.
XX
XX Sequence 8 AA;
SQ
Query Match 43.2%; Score 35; DB 22; Length 8;
Best Local Similarity 50.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 4 PPMWSPVW 11
DB 1 PPSWDQMW 8
RESULT 8
AAJ01682
ID AAJ01682 standard; Peptide; 8 AA.
XX AC
XX AAJ01682;
XX 02-JUL-2001 (first entry)
XX Hepatitis C virus epitope #1673.
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX Hepatitis C virus.
OS
XX W02C0121:189-Al.
PN 29-MAR-2001.
XX 19-JUL-2000; 20CWO-US19774.
XX 19-JUL-1999; 99US-0357737.
XX (EPIM-) EPIMMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
XX A new composition useful as a vaccine against hepatitis C virus -
PS Disclosure; Page 142; 214pp; English.
CC The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention.
XX
XX Sequence 8 AA;
SQ
Query Match 43.2%; Score 35; DB 22; Length 8;
Best Local Similarity 50.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 4 PPMWSPVW 11
DB 1 PPSWDQMW 8
RESULT 9
AAJ01136
ID AAJ01136 standard; Peptide; 9 AA.
XX AC
XX AAJ01136;

```

XX 02-JUL-2001 (first entry)
XX Hepatitis C virus epitope #1127.
DE Hepatitis C virus epitope #1127.
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.
XX Hepatitis C virus.
OS Hepatitis C virus.
XX WO200121189-A1.
PN 29-MAR-2001.
PD 19-JUL-2000; 2000WO-US19774.
XX 19-JUL-1999; 99US-0357737.
PR (EPIM-) EPIMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;
XX WPI; 2001-308046/32.
DR A new composition useful as a vaccines against hepatitis C virus -
PT Disclosure; Page 128; 214pp; English.
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention.
XX Sequence 9 AA;
SQ
Query Match 43.2%; Score 35; DB 22; Length 9;
Best Local Similarity 50.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

XX 4 PPMWSPVW 11
DB ||| :|
2 PPSWDQMW 9

RESULT 11
AAJ01801
ID AAJ01801 standard; Peptide; 9 AA.
XX AC AAJ01801;
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #1792.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX KW antiviral.
XX OS Hepatitis C virus.
XX PN WO200121189-A1.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US19774.
XX PR 19-JUL-1999; 99US-0357737.
XX PA (EPIM-) EPIMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;

Query Match 43.2%; Score 35; DB 22; Length 9;
Best Local Similarity 50.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

XX 4 PPMWSPVW 11
DB ||| :|
2 PPSWDQMW 9

RESULT 10
AAJ01801
ID AAJ01801 standard; Peptide; 9 AA.
XX AC AAJ01801;
XX DT 02-JUL-2001 (first entry)
XX DE Hepatitis C virus epitope #1792.
XX KW Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX KW antiviral.
XX OS Hepatitis C virus.
XX PN WO200121189-A1.
XX PD 29-MAR-2001.
XX PF 19-JUL-2000; 2000WO-US19774.
XX PR 19-JUL-1999; 99US-0357737.
XX PA (EPIM-) EPIMUNE INC.
XX Sette A, Sidney J, Southwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;

```

```

XX WPI; 2001-308046/32.
XX A new composition useful as a vaccines against hepatitis C virus -
XX Disclosure; Page 145; 214pp; English.
XX The present invention describes a composition comprising a prepared
CC hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention.
XX Sequence 9 AA;
SQ
Query Match 43.2%; Score 35; DB 22; Length 9;
Best Local Similarity 50.0%; Pred. No. 9.3e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

XX 4 PPMWSPVW 11
DB ||| :|
2 PPSWDQMW 9

RESULT 11
AAW24284
ID AAW24284 standard; peptide; 10 AA.
XX AC AAW24284;
XX DT 07-MAY-1998 (first entry)
XX DE CBP1/RBPJk interaction domain of EBNA2.
XX KW CBP1/RBPJk interaction domain; EBNA2; transplant.
XX OS Epstein-Barr Virus.
XX PN WO9730081-A1.
XX PD 21-AUG-1997.
XX PF 12-FEB-1997; 97WO-US02243.
XX PR 13-FEB-1996; 96US-0011557.
XX PA (UYCO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.
XX PI Hayward SD, Ling PD;
XX WPI; 1997-424974/39.
XX PT Peptide with homology to CBP1-RBPJk interaction domain of EBNA2 -
XX useful to provide protection against Epstein-Barr virus
XX Claim 12; Page 46; 62pp; English.
XX AAW24280-W24284 represent five new peptides which display substantial
CC homology with the CBP1/RBPJk interaction domain of EBNA2. In particular,
CC the peptides include amino acids 318-327 of EBNA2 (i.e. conserved
CC region 6, or CR6). The peptides, when administered in an amount
CC sufficient to inhibit interaction of EBV EBNA2 with CBP1, are used to
CC protect mammals prior to or following a transplant, against Epstein-Barr
CC virus. Labelled peptides are used to detect factors that interact with
CC EBV EBNA2.
XX Sequence 10 AA;
SQ
Query Match 43.2%; Score 35; DB 18; Length 10;
Best Local Similarity 60.0%; Pred. No. 95;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

XX 1 THRPWMSPV 10

```


SQ Sequence 10 AA;

Query Match 43.2%; Score 35; DB 22; Length 10;
Best Local Similarity 50.0%; Pred. No. 95;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 PPMWSPVW 11
||| :
Db 3 PPSWDQW 10

RESULT 15

AAJ01581
ID AAJ01581 standard; Peptide; 11 AA.

XX AAJ01581;

AC AAJ01581;

XX 02-JUL-2001 (first entry)

XX Hepatitis C virus epitope #1572.

DE Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
XX Hepatitis C virus; HCV; epitope; vaccine; immunogen; HLA-binding motif;
KW antiviral.

XX Hepatitis C virus.

XX WO200121189-A1.

XX 29-MAR-2001.

XX 19-JUL-2000; 2000WO-0519774.

XX 19-JUL-1999; 99US-0357737.

XX (EPIM-) EPIWONE INC.

XX Sette A, Sidney J, Scottwood S, Livingston BD, Chesnut R;
PI Baker DM, Celis E, Kubo RT, Grey HM;

XX WPI; 2001-308046/32.

XX A new composition useful as a vaccines against hepatitis C virus -

XX Disclosure; Page 140; 2:4pp; English.

XX The present invention describes a composition comprising a prepared
CC Hepatitis C virus (HCV) epitope such as those given in AAJ00010-AAJ04121.
CC These are derived from HCV HLA-binding motifs. They are useful in
CC vaccines for the prevention and treatment of HCV infection in humans. The
CC present sequence is an epitope used in the disclosure of the invention.

SQ Sequence 11 AA;

Query Match 43.2%; Score 35; DB 22; Length 11;
Best Local Similarity 50.0%; Pred. No. 1e+02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 PPMWSPVW 11
||| :
Db 4 PPSWDQW 11

Search completed: October 22, 2003, 20:51:59
Job time : 84 secs

GenCore version: 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 22, 2003, 20:55:02 ; Search time 67 Seconds
(without alignments)
29.993 Million cell updates/sec

Title: us-09-995-804A-2
Perfect score: 81
Sequence: 1 *HPPWSPWP 12

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 629382 seqs, 167460630 residues

Total number of hits satisfying chosen parameters: 92044

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA: *

1:	/cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/prodata/2/pubpaa/PCR_NEW_PUB.pep.*
3:	/cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/prodata/2/pubpaa/PCRUS_PUBCOMB.pep.*
7:	/cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
10:	/cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
11:	/cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
12:	/cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
14:	/cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
15:	/cgn2_6/prodata/2/pubpaa/US10_PUBCOMB.pep.*
16:	/cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	81	100.0	12	10	US-09-995-804A-2
2	35	43.2	9	7	US-08-344-824-93
3	35	43.2	10	7	US-08-344-824-193
4	35	43.2	10	9	US-09-739-852-7
5	35	43.2	12	15	US-10-075-869-45
6	33	40.7	10	9	US-09-739-852-6
7	33	40.7	10	15	US-10-226-007-1299
8	33	40.7	11	15	US-10-226-007-1300
9	33	40.7	11	15	US-10-226-007-1313
10	33	40.7	12	12	US-10-286-457-48
11	33	40.7	12	15	US-10-226-007-1301
12	33	40.7	12	15	US-10-226-007-1314
13	33	40.7	12	15	US-10-226-007-1327
14	31	38.3	10	12	US-10-161-791-141
15	30	37.0	6	9	US-09-739-852-1

15	30	37.0	8	15	US-10-226-007-1453	Sequence 1453, Ap
17	30	37.0	9	12	US-10-062-109A-116	Sequence 116, App
18	30	37.0	9	12	US-10-005-480A-116	Sequence 116, App
19	30	37.0	9	15	US-10-226-007-1455	Sequence 1455, Ap
20	30	37.0	10	10	US-09-910-552-47	Sequence 47, Appl
21	30	37.0	10	10	US-09-910-552-64	Sequence 64, Appl
22	30	37.0	10	10	US-09-910-552-66	Sequence 66, Appl
23	30	37.0	10	12	US-10-062-109A-163	Sequence 163, App
24	30	37.0	10	12	US-10-062-109A-218	Sequence 218, App
25	30	37.0	10	12	US-10-062-109A-624	Sequence 624, App
26	30	37.0	10	12	US-10-005-480A-163	Sequence 163, App
27	30	37.0	10	12	US-10-005-480A-218	Sequence 218, App
28	30	37.0	10	12	US-10-005-480A-624	Sequence 624, App
29	30	37.0	10	15	US-10-226-007-1458	Sequence 1458, Ap
30	30	37.0	11	9	US-09-739-852-2	Sequence 2, Appl
31	30	37.0	11	15	US-10-226-007-1461	Sequence 1461, Ap
32	30	37.0	12	15	US-10-226-007-1464	Sequence 1464, Ap
33	30	37.0	12	15	US-10-254-446A-170	Sequence 170, App
34	30	37.0	12	15	US-10-254-446A-245	Sequence 245, App
35	29	35.8	8	15	US-10-226-007-1297	Sequence 1297, Ap
36	29	35.8	8	15	US-10-226-007-1310	Sequence 1310, Ap
37	29	35.8	9	12	US-09-793-451-527	Sequence 527, App
38	29	35.8	9	12	US-10-283-722-527	Sequence 527, App
39	29	35.8	9	15	US-10-226-007-1298	Sequence 1298, Ap
40	29	35.8	9	15	US-10-226-007-1311	Sequence 1311, Ap
41	29	35.8	9	15	US-10-226-007-1324	Sequence 1324, Ap
42	29	35.8	10	12	US-09-572-404B-2240	Sequence 2240, Ap
43	29	35.8	10	12	US-09-793-451-578	Sequence 578, App
44	29	35.8	10	12	US-10-283-722-578	Sequence 578, App
45	29	35.8	10	15	US-10-226-007-1312	Sequence 1312, Ap

ALIGNMENTS

RESULT 1
US-09-995-804A-2
; Sequence 2, Application US/09995804A
; Patent No. US20020115824A1
; GENERAL INFORMATION:
; APPLICANT: Engler, Jeffrey A
; Lee, Jae Hwy
; Colliawan, James F
; Moore, Bryan A
; TITLE OF INVENTION: Receptor-Mediated Uptake of Peptides
; that Bind the Human Transferrin Receptor
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hendricks and Assoc.
; STREET: P.O. Box 2509
; CITY: Fairfax
; STATE: VA
; COUNTRY: U.S.A
; ZIP: 22031
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/995,804A
; FILING DATE: 29-No. US20020115824A1-2001
; CLASSIFICATION: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Hendricks, Glenra M
; REGISTRATION NUMBER: 32,535
; REFERENCE/DOCKET NUMBER: engler1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703/425-8405
; TELEFAX: 703/425-8406
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids

Sub p2

```
;
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-995-804A-2

Query Match 100.0%; Score 81; DB 10; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.00024;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 THRPFMSPVWP 12
Db 1 THRPFMSPVWP 12

RESULT 2
US-08-344-924-93
; Sequence 93, Application US/08344824
; Publication No. US20030152580A1
; GENERAL INFORMATION:
; APPLICANT: SETTE, Alessandro
; TITLE OF INVENTION: H2A BINDING PEPTIDES AND THEIR USES
; NUMBER OF SEQUENCES: 399
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Stewart Street Tower, 20th
; STREET: Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/344,924
; FILING DATE: 23-NOV-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/278,634
; FILING DATE: 21-JUL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Bastian, Kevin L.
; REGISTRATION NUMBER: 34,774
; REFERENCE/DOCKET NUMBER: 14137-80-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 543-9600
; TELEFAX: (415) 543-5043
; INFORMATION FOR SEQ ID NO: 93:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-344-824-193

Query Match 43.2%; Score 35; DB 7; Length 10;
Best Local Similarity 50.0%; Pred. No. 1.7e-02;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 PPWMSPVW 11
Db 3 PPSWDQMW 10

RESULT 4
US-09-739-852-7
; Sequence 7, Application US/09739852
; Patent No. US20010014337A1
; GENERAL INFORMATION:
; APPLICANT: Hayward, Diane S.
; APPLICANT: Ling, Paul D.
; TITLE OF INVENTION: EBNA2 PEPTIDES AND METHODS OF USING SAME
; FILE REFERENCE: 87512
; CURRENT APPLICATION NUMBER: US/09/739,852
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 09/133,341
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
```

OTHER INFORMATION: Description of Unknown Virus Organism: Artificial
US-09-739-852-7

Query Match 43.2%; Score 35; DB 9; Length 10;
Best Local Similarity 60.0%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 THRPWKSPV 10
DB 1 TPGPPWPEV 10

RESULT 5
US-10-075-869-45
; Sequence 45, Application US/10075869
; Publication No. US20030104622A1
; GENERAL INFORMATION:
; APPLICANT: Robbins, Paul D.
; APPLICANT: Mi, Zhibao
; APPLICANT: Friizell, Raymond
; APPLICANT: Glorioso, Joseph C.
; APPLICANT: Gambotto, Andrea
; TITLE OF INVENTION: IDENTIFICATION OF PEPTIDES THAT
; TITLE OF INVENTION: FACILITATE UPTAKE AND CYTOPLASMIC AND/OR NUCLEAR TRANSPORT
; TITLE OF INVENTION: OF PROTEINS, DNA AND VIRUSES
; FILE REFERENCE: AP12573-AAA 072396.0237
; CURRENT APPLICATION NUMBER: US/10/075,869
; CURRENT FILING DATE: 2002-02-13
; PRIOR APPLICATION NUMBER: 60/151,980
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: 60/188,944
; PRIOR FILING DATE: 2000-03-13
; NUMBER OF SEQ ID NOS: 99
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 45
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: random peptide library
US-10-075-869-45

Query Match 43.2%; Score 35; DB 15; Length 12;
Best Local Similarity 66.7%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 7 WSPWVP 12
DB 4 WLPLWP 9

RESULT 6
US-09-739-852-6
; Sequence 6, Application US/09739852
; Patent No. US20010014337A1
; GENERAL INFORMATION:
; APPLICANT: Hayward, Diane S.
; APPLICANT: Ling, Paul D.
; TITLE OF INVENTION: EBNA2 PEPTIDES AND METHODS OF USING SAME
; FILE REFERENCE: 875.12
; CURRENT APPLICATION NUMBER: US/09/739,852
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 09/133,341
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Virus Organism: Artificial
US-09-739-852-6

Query Match 40.7%; Score 33; DB 9; Length 10;
Best Local Similarity 57.1%; Pred. No. 3.1e+02;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PPMWSPV 10
DB 4 PPMWPPI 10

RESULT 7
US-10-226-007-1299
; Sequence 1299, Application US/10226007
; Publication No. US20030105277A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226,007
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1299
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Human herpesvirus 7
US-10-226-007-1299

Query Match 40.7%; Score 33; DB 15; Length 10;
Best Local Similarity 55.6%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 PPMWSPWVP 12
DB 2 PPWYPSMP 10

RESULT 8
US-10-226-007-1300
; Sequence 1300, Application US/10226007
; Publication No. US20030105277A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kenton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226,007
; CURRENT FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1300
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Human herpesvirus 7
US-10-226-007-1300

Query Match 40.7%; Score 33; DB 15; Length 11;
Best Local Similarity 55.6%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 4 PPMWSPWVP 12
DB 2 PPWYPSMP 10

```
RESULT 9
; US-10-226-007-1313
; Sequence 1313, Application US/10226007
; Publication No. US20030105277A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kerton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226.007
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1313
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Human herpesvirus 7
; US-10-226-007-1313

Query Match 40.7%; Score 33; DB 15; Length 11;
Best Local Similarity 55.6%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 PPMWSPVWP 12
Db 3 PPYWPSMP 11

RESULT 10
; US-10-286-457-48
; Sequence 48, Application US/10286457
; Publication No. US20030166004A1
; GENERAL INFORMATION:
; APPLICANT: JENO GVURIS et al.
; TITLE OF INVENTION: ENDOTHELIAL-CELL BINDING PEPTIDES FOR DIAGNOSIS AND THERAPY
; FILE REFERENCE: GPCI-P01-178
; CURRENT APPLICATION NUMBER: US/10/286.457
; PRIOR FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 66/334822
; PRIOR FILING DATE: 2001-11-01
; NUMBER OF SEQ ID NOS: 684
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: artificial sequence isolated from random peptide libraries, based
; OTHER INFORMATION: ability to selectively bind to endothelial cells
; US-10-286-457-48

Query Match 40.7%; Score 33; DB 12; Length 12;
Best Local Similarity 45.5%; Pred. No. 3.6e+02;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 2 HRPWMSPVWP 12
Db 2 HEPGAWLPLSP 12

RESULT 11
; US-10-226-007-1301
; Sequence 1301, Application US/10226007
; Publication No. US20030105277A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kerton
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226.007
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1301
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Human herpesvirus 7
; US-10-226-007-1301

Query Match 40.7%; Score 33; DB 15; Length 12;
Best Local Similarity 55.6%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 PPMWSPVWP 12
Db 3 PPYWPSMP 11

RESULT 12
; US-10-226-007-1314
; Sequence 1314, Application US/10226007
; Publication No. US20030105277A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kerton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226.007
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1314
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Human herpesvirus 7
; US-10-226-007-1314

Query Match 40.7%; Score 33; DB 15; Length 12;
Best Local Similarity 55.6%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 4 PPMWSPVWP 12
Db 3 PPYWPSMP 11

RESULT 13
; US-10-226-007-1327
; Sequence 1327, Application US/10226007
; Publication No. US20030105277A1
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Inc.
; APPLICANT: Morham, Scott
; APPLICANT: Zavitz, Kerton
; APPLICANT: Hobden, Adrian
; TITLE OF INVENTION: Compositions and Therapeutic Methods for Viral Infection
; FILE REFERENCE: 5005.01
; CURRENT APPLICATION NUMBER: US/10/226.007
; PRIOR FILING DATE: 2002-11-15
; PRIOR APPLICATION NUMBER: US 60/313,883
; PRIOR FILING DATE: 2001-08-21
; NUMBER OF SEQ ID NOS: 1673
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1327
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Human herpesvirus 7
; US-10-226-007-1327
```


/ SEQ ID NO 1327
/ LENGTH: 12
/ TYPE: PRT
/ ORGANISM: Human herpesvirus 7
US-10-226-007-1327

Query Match 40.7%; Score 33; DB 15; Length 12;
Best Local Similarity 55.6%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 4; Indels 0;

QY 4 PPMWSPVMP 12
|||
Db 4 PPWYPSMP 12

RESULT 14
US-10-161-791-141
/ Sequence 141, Application US/10:6179;
/ Publication No. US20030196963A1
/ GENERAL INFORMATION:
/ APPLICANT: SPARKS, Andrew B.
/ APPLICANT: KAY, Brian K.
/ APPLICANT: THORN, Judith M.
/ APPLICANT: QUILLIAM, Lawrence A.
/ APPLICANT: DER, Channing J.
/ APPLICANT: FOWLKES, Dana M.
/ APPLICANT: RIDER, James E.
/ TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
/ TITLE OF INVENTION: ISOLATING AND USING SAME
/ NUMBER OF SEQUENCES: 467
/ CORRESPONDENCE ADDRESS:
/ ADDRESSES: Pennie & Edmonds
/ STREET: 1155 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036-2711
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent'n Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/161,791;
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/08/602,999
/ FILING DATE: 16-FEB-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Misrock, S. Leslie
/ REGISTRATION NUMBER: 18,872
/ REFERENCE/DOCKET NUMBER: 1101-202
/ TELEPHONE: (212) 790-909C
/ TELEFAX: (212) 869-9741/8864
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 141:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 10 amino acids
/ TYPE: amino acid
/ TOPOLOGY: unknown
/ MOLECULE TYPE: peptide
/ FEATURE:
/ NAME/KEY: Modified-site
/ LOCATION: 9
/ OTHER INFORMATION: /note= "X = a hydrophobic amino
/ OTHER INFORMATION: acid residue"
US-10-161-791-141

Query Match 38.3%; Score 31; DB 12; Length 10;
Best Local Similarity 66.7%; Pred. No. 5.6e+02;
Matches 4; Conservative 0; Mismatches 2; Indels 0;

QY 4 PPMWSP 9
|||
Db 1 PPXWXP 6

RESULT 15
US-09-739-852-1
/ Sequence 1, Application US/09739852
/ Patent No. US20010014337A1
/ GENERAL INFORMATION:
/ APPLICANT: Hayward, Diane S.
/ APPLICANT: Ling, Paul D.
/ TITLE OF INVENTION: BENA2 PEPTIDES AND METHODS OF USING SAME
/ FILE REFERENCE: 87512
/ CURRENT APPLICATION NUMBER: US/09/739,852
/ CURRENT FILING DATE: 2000-12-18
/ PRIOR APPLICATION NUMBER: 09/133,341
/ PRIOR FILING DATE: 2000-12-18
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: Patent'n Ver. 2.0
/ SEQ ID NO 1
/ LENGTH: 6
/ TYPE: PRT
/ ORGANISM: Epstein-Barr Virus
US-09-739-852-1

Query Match 37.0%; Score 30; DB 9; Length 6;
Best Local Similarity 66.7%; Pred. No. 5.6e+05;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 PPMWSP 9
|||
Db 1 PPWPPP 6

Search completed: October 22, 2003, 21:03:52
Job time : 68 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OX protein - protein search, using sw model

Run on: October 22, 2003, 20:52:32 ; Search time 28 Seconds
(without alignments)
18.133 Million cell updates/sec

Title: US-09-995-804A-2

Perfect score: 81

Sequence: 1 THRPMPWSPVWP 12

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 105880

Minimum DB seq length: 0
Maximum DB seq length: 12

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PTJUS_COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	43.2	6	3	US-08-444-818-445
2	35	43.2	10	3	US-09-133-341-7
3	35	43.2	10	4	US-09-739-852-7
4	33	40.7	10	3	US-09-133-341-6
5	33	40.7	10	4	US-09-739-852-6
6	33	40.7	12	1	US-08-190-788A-101
7	33	40.7	12	1	US-08-383-474B-106
8	33	40.7	12	1	US-08-465-391A-101
9	33	40.7	12	2	US-08-464-538B-101
10	33	40.7	12	2	US-08-463-076E-147
11	32	39.5	9	2	US-08-318-856A-67
12	31	38.3	10	3	US-08-602-999A-141
13	31	38.3	10	4	US-08-500-124-141
14	30.5	37.7	12	1	US-08-014-426-32
15	30.5	37.7	12	1	US-08-014-426-47
16	30.5	37.7	12	5	PCT-US94-01319-32
17	30.5	37.7	12	5	PCT-US94-01319-47
18	30	37.0	6	3	US-09-133-341-1
19	30	37.0	6	4	US-09-739-852-1
20	30	37.0	9	3	US-08-336-553A-14
21	30	37.0	9	4	US-08-439-157-14
22	30	37.0	10	4	US-09-437-895-14
23	30	37.0	10	3	US-08-925-002-47
24	30	37.0	10	3	US-08-925-002-64
25	30	37.0	10	3	US-08-925-002-66
26	30	37.0	11	1	US-08-175-471-3
27	30	37.0	11	2	US-08-429-054A-3

28	30	37.0	11	2	US-08-718-777-3	Sequence 3, Appli
29	30	37.0	11	3	US-09-051-341-3	Sequence 3, Appli
30	30	37.0	11	3	US-09-133-341-2	Sequence 2, Appli
31	30	37.0	11	4	US-09-739-852-2	Sequence 2, Appli
32	29	35.8	9	1	US-08-227-184A-3	Sequence 3, Appli
33	29	35.8	11	1	US-08-465-391A-326	Sequence 326, App
34	29	35.8	11	2	US-08-526-764-1	Sequence 1, Appli
35	29	35.8	11	2	US-08-464-538B-324	Sequence 324, App
36	29	35.8	11	2	US-08-463-076E-40	Sequence 40, Appl
37	28.5	35.2	8	4	US-09-007-288E-70	Sequence 70, Appl
38	28.5	35.2	8	4	US-09-007-288E-73	Sequence 73, Appl
39	28.5	35.2	8	4	US-09-007-288E-75	Sequence 75, Appl
40	28	34.6	5	1	US-07-789-184-86	Sequence 88, Appl
41	28	34.6	5	1	US-08-475-263-88	Sequence 88, App
42	28	34.6	5	1	US-08-485-886-88	Sequence 88, Appl
43	28	34.6	5	2	US-08-477-362-88	Sequence 88, App
44	28	34.6	5	2	US-08-477-134-88	Sequence 88, App
45	28	34.6	5	3	US-08-473-489A-88	Sequence 88, Appl

ALIGNMENTS

RESULT 1
US-08-444-818-445
; Sequence 445, Application US/C8444818
; Patent No. 6150067
; GENERAL INFORMATION:
; APPLICANT: Chien, David Y.
; APPLICANT: Rutter, William J.
; TITLE OF INVENTION: NANV Diagnostics and Vaccines
; NUMBER OF SEQUENCES: 777
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: CA
; COUNTRY: USA
; ZIP: 94608-2916
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,518
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/403,590
; FILING DATE: 14-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Harbin, Alisa A.
; REGISTRATION NUMBER: 33,895
; REFERENCE/DOCKET NUMBER: 0110.002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (508)359-3876
; TELEFAX: (508)359-3885
; INFORMATION FOR SEQ ID NO: 445:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-444-818-445

Query Match 43.2%; Score 35; DB 3; Length 8;
Best Local Similarity 50.0%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 4 PPMWSPVW 11
||| :|

```
Db      1 2PSWDQNMW 8

RESULT 2
US-09-133-341-7
; Sequence 7, Application US/09:33341A
; Patent No. 6162440
; GENERAL INFORMATION:
; APPLICANT: Hayward, Diane S.
; APPLICANT: Ling, Paul D.
; TITLE OF INVENTION: EBNA2 PEPTIDES AND METHODS OF USING SAME
; FILE REFERENCE: 87512
; CURRENT APPLICATION NUMBER: US/09/133,341A
; CURRENT FILING DATE: 1998-08-13
; EARLIER APPLICATION NUMBER: WO 97US2243
; EARLIER FILING DATE: 1997-02-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Virus Organism: Artificial
US-09-133-341-7

Query Match: 43.2%; Score 35; DB 3; Length 10;
Best Local Similarity 60.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 THRPWMSPV 10
       |||||
Db      1 TPGPPWPPV 10

RESULT 3
US-09-739-852-7
; Sequence 7, Application US/09739852
; Patent No. 6495144
; GENERAL INFORMATION:
; APPLICANT: Hayward, Diane S.
; APPLICANT: Ling, Paul D.
; TITLE OF INVENTION: EBNA2 PEPTIDES AND METHODS OF USING SAME
; FILE REFERENCE: 87512
; CURRENT APPLICATION NUMBER: US/09/739,852
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 09/133,341
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Virus Organism: Artificial
US-09-739-852-7

Query Match: 43.2%; Score 35; DB 4; Length 10;
Best Local Similarity 60.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY      1 THRPWMSPV 10
       |||||
Db      1 TPGPPWPPV 10

RESULT 4
US-09-133-341-6
; Sequence 6, Application US/09:33341A
; Patent No. 6162440
; GENERAL INFORMATION:
; APPLICANT: Hayward, Diane S.
; APPLICANT: Ling, Paul D.
; TITLE OF INVENTION: EBNA2 PEPTIDES AND METHODS OF USING SAME
; FILE REFERENCE: 87512
; CURRENT APPLICATION NUMBER: US/09/133,341A
; CURRENT FILING DATE: 1998-08-13
; EARLIER APPLICATION NUMBER: WO 97US2243
; EARLIER FILING DATE: 1997-02-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Virus Organism: Artificial
US-09-133-341-6

Query Match: 40.7%; Score 33; DB 3; Length 10;
Best Local Similarity 57.1%; Pred. No. 52;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      4 PPMWSPV 10
       |||||
Db      4 PPMWPPV 10

RESULT 5
US-09-739-852-6
; Sequence 6, Application US/09739852
; Patent No. 6495144
; GENERAL INFORMATION:
; APPLICANT: Hayward, Diane S.
; APPLICANT: Ling, Paul D.
; TITLE OF INVENTION: EBNA2 PEPTIDES AND METHODS OF USING SAME
; FILE REFERENCE: 87512
; CURRENT APPLICATION NUMBER: US/09/739,852
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: 09/133,341
; PRIOR FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Virus Organism: Artificial
US-09-739-852-6

Query Match: 40.7%; Score 33; DB 4; Length 10;
Best Local Similarity 57.1%; Pred. No. 52;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      4 PPMWSPV 10
       |||||
Db      4 PPMWPPV 10

RESULT 6
US-08-190-788A-101
; Sequence 101, Application US/08190788A
; Patent No. 5608035
; GENERAL INFORMATION:
; APPLICANT: Yanofsky, Stephen D.
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Baldwin, David N.
; APPLICANT: Jacobs, Jeff W.
; TITLE OF INVENTION: Peptides and Compounds That Bind to the
; TITLE OF INVENTION: IL-1 Receptor
; NUMBER OF SEQUENCES: 312
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Affymax Technologies N.V.
; STREET: 4001 Miranda Avenue
```

CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/90,788A
FILING DATE: 02-FEB-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/847,567
FILING DATE: 05-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Stevens, Lauren L.
REGISTRATION NUMBER: 36,691
REFERENCE/DOCKET NUMBER: 1019.1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-496-2300
TELEFAX: 415-424-0832
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-190-788A-101

Query Match 40.7%; Score 33; DB 1; Length 12;
Best Local Similarity 60.0%; Pred. No. 63;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WSPW 11
Db 8 WAPIW 12

RESULT 7
US-08-383-474B-106
Sequence 106, Application US/08383474B
Patent No. 5767234
GENERAL INFORMATION:
APPLICANT: Yanofsky, Stephen D.
APPLICANT: Barrett, Ronald W.
APPLICANT: Baldwin, David N.
APPLICANT: Jacobs, Jeff W.
TITLE OF INVENTION: Peptides and Compounds That Bind to
TITLE OF INVENTION: the IL-1 Receptor
NUMBER OF SEQUENCES: 314
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend & Townsend & Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/383,474B
FILING DATE: 01-FEB-1995
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,788
FILING DATE: 02-FEB-1994
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:
NAME: Stevens, Lauren L.
REGISTRATION NUMBER: 36,691
REFERENCE/DOCKET NUMBER: 1019.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-496-2300
TELEFAX: 415-424-0832
INFORMATION FOR SEQ ID NO: 106:
SEQUENCE CHARACTERISTICS:
LENGTH: 12 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-383-474B-106

Query Match 40.7%; Score 33; DB 1; Length 12;
Best Local Similarity 60.0%; Pred. No. 63;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WSPW 11
Db 8 WAPIW 12

RESULT 8
US-08-465-391A-101
Sequence 101, Application US/08465391A
Patent No. 5786331
GENERAL INFORMATION:
APPLICANT: Barrett, Ronald W.
APPLICANT: Yanofsky, Stephen D.
APPLICANT: Baldwin, David
APPLICANT: Jacobs, Jeff W.
APPLICANT: Bovy, Philippe R.
APPLICANT: Leahy, Ellen M.
APPLICANT: Pottorf, Richard S.
TITLE OF INVENTION: Peptides and Compounds That Bind to the
TITLE OF INVENTION: IL-1 Receptor
NUMBER OF SEQUENCES: 405
CORRESPONDENCE ADDRESS:
ADDRESS: Townsend and Townsend and Crew
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,391A
FILING DATE: 05-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/373,474
FILING DATE: 01-FEB-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/190,788
FILING DATE: 02-FEB-1994
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: No. 5786331viel, Vern
REGISTRATION NUMBER: 32,483
REFERENCE/DOCKET NUMBER: 16528A-001840/1019.2A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 101:
SEQUENCE CHARACTERISTICS:

; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-465-391A-101

Query Match 40.7%; Score 33; DB 1; Length 12;
Best Local Similarity 60.0%; Pred. No. 63;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WSPVW 11
|:|:
Db 8 WAPIW 12

RESULT 9
US-08-464-538B-101
; Sequence 101, Application US/58464538B
; Patent No. 5861476
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Yanofsky, Stephen D.
; APPLICANT: Baldwin, David
; APPLICANT: Jacobs, Jeff W.
; APPLICANT: Bovy, Phillippe R.
; APPLICANT: Leary, Ellen M.
; APPLICANT: Portoff, Richard S.
; TITLE OF INVENTION: Peptides and Compounds That Bind to the
; TITLE OF INVENTION: IL-1 Receptor
; NUMBER OF SEQUENCES: 402
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, 8th Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/464,538B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/373,474
; FILING DATE: 01-FEB-1995
; CLASSIFICATION: 514
; INFORMATION FOR SEQ ID NO: 101:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-464-538B-101

Query Match 40.7%; Score 33; DB 2; Length 12;
Best Local Similarity 60.0%; Pred. No. 63;

Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
Qy 7 WSPVW 11
|:|:
Db 8 WAPIW 12

RESULT 10
US-08-463-076E-147
; Sequence 147, Application US/08463076E
; Patent No. 5880096
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Yanofsky, Stephen D.
; TITLE OF INVENTION: Peptides and Compounds That Bind to the
; TITLE OF INVENTION: IL-1 Receptor
; NUMBER OF SEQUENCES: 392
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/463,076E
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Snyder, Joseph R.
; REGISTRATION NUMBER: 39,381
; REFERENCE/DOCKET NUMBER: 16528A-001850US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 147:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-463-076E-147

Query Match 40.7%; Score 33; DB 2; Length 12;
Best Local Similarity 60.0%; Pred. No. 63;
Matches 3; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 7 WSPVW 11
|:|:
Db 8 WAPIW 12

RESULT 11
US-08-318-856A-67
; Sequence 67, Application US/08318856A
; Patent No. 5972351
; GENERAL INFORMATION:
; APPLICANT: Adrian V.S. Hill, et al.
; TITLE OF INVENTION: PLASMODIUM FALCIPARUM MHC CLASS I-
; TITLE OF INVENTION: RESTRICTED CTL EPITOPES DERIVED FROM PRE-ERYTHROCYTIC STAGE
; TITLE OF INVENTION: ANTIGENS (AS AMENDED)
; NUMBER OF SEQUENCES: 86
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wenderoth, Lind & Ponack, J.L.P.
; STREET: 2033 K Street, N.W., Suite 800
; CITY: Washington
; STATE: D.C.

```

; COUNTRY: U.S.A.
; ZIP: 20006
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk, 3.5 inch, 1.44 mb
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WordPerfect 5.1+
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/318,956A
; FILING DATE: October 3, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 08 068.8
; FILING DATE: April 3, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 92 17 704.7
; FILING DATE: August 20, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/GB93/00711
; FILING DATE: April 5, 1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lee Cheng
; REGISTRATION NUMBER: 40,949
; REFERENCE/DOCKET NUMBER: 263-PP1R1577US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 721-8200
; TELEFAX: (202) 721-8250
; INFORMATION FOR SEQ ID NO: 67:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9 amino acid residues
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Peptide
; US-08-318-856A-67

```

```

Query Match 39.5%; Score 32; DB 2; Length 9;
Best Local Similarity 57.1%; Pred. No. 2.5e+05;
Matches 4; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

```

```

CY 4 PPMWSPV 10
DB 1 PPKWEPL 7

```

```

RESULT 12
; Sequence 141, Application: US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,999A

```

```

; FILING DATE: 16-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 141:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 9
; OTHER INFORMATION: /note= "X = a hydrophobic amino
; acid residue"
; US-08-602-999A-141

```

```

Query Match 38.3%; Score 31; DB 3; Length 10;
Best Local Similarity 66.7%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

CY 4 PPMWSP 9
DB 1 PPKWXP 6

```

```

RESULT 13
; Sequence 141, Application US/09500124
; Patent No. 6432920
; GENERAL INFORMATION:
; APPLICANT: SPARKS, Andrew B.
; APPLICANT: KAY, Brian K.
; APPLICANT: THORN, Judith M.
; APPLICANT: QUILLIAM, Lawrence A.
; APPLICANT: DER, Channing J.
; APPLICANT: FOWKES, Dana M.
; APPLICANT: RIDER, James E.
; TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
; ISOLATING AND USING SAME
; NUMBER OF SEQUENCES: 467
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/500,124
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/602,999
; FILING DATE: 16-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Misrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864

```

```
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 141:
SEQUENCE CHARACTERISTICS:
LENGTH: 10 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 9
OTHER INFORMATION: /note="X" = a hydrophobic amino
OTHER INFORMATION: acid residue"
US-09-500-124-141

Query Match      38.3%; Score 31; DB 4; Length 10;
Best Local Similarity 66.7%; Pred. No. 99;
Matches 4; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      4 PPMWSP 9
DB      : PPMWXP 6

RESULT 14
US-08-014-426-32
; Sequence 32, Application US/08014426
; Patent No. 5512435
; GENERAL INFORMATION:
; APPLICANT: Renschler, Markus F.
; APPLICANT: Levy, Ronald
; APPLICANT: Bhatt, Ramesh
; APPLICANT: Dower, William
; TITLE OF INVENTION: RECEPTOR-BINDING ANTIPROLIFERATIVE
; TITLE OF INVENTION: PEPTIDES
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.23
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,426
; FILING DATE: 05-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 5490A-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2422
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
US-08-014-426-32

Query Match      37.7%; Score 30.5; DB 1; Length 12;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY      4 PPMW-SPWV 11
DB      : PPMWQSRVW 11
```

Search completed: October 22, 2003, 21:02:06
Job time : 29 secs

```
DB      3 PPMWQSRVW 11

RESULT 15
US-08-014-426-47
; Sequence 47, Application US/08014426
; Patent No. 5512435
; GENERAL INFORMATION:
; APPLICANT: Renschler, Markus F.
; APPLICANT: Levy, Ronald
; APPLICANT: Bhatt, Ramesh
; APPLICANT: Dower, William
; TITLE OF INVENTION: RECEPTOR-BINDING ANTIPROLIFERATIVE
; TITLE OF INVENTION: PEPTIDES
; NUMBER OF SEQUENCES: 65
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: 379 Lytton Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: US
; ZIP: 94301
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,426
; FILING DATE: 05-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 5490A-204
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 326-2400
; TELEFAX: (415) 326-2422
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
US-08-014-426-47

Query Match      37.7%; Score 30.5; DB 1; Length 12;
Best Local Similarity 66.7%; Pred. No. 1.4e+02;
Matches 6; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY      4 PPMW-SPWV 11
DB      : PPMWQSRVW 11
```